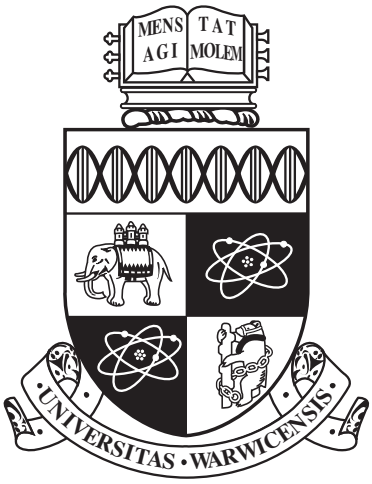



# The Combinatorics of the Cluster and Virial Expansions

by

**Stephen James Tate**

**Thesis**

Submitted to the University of Warwick

for the degree of

**Doctor of Philosophy**

**Department of Mathematics**

March 2014

THE UNIVERSITY OF
WARWICK

# Contents

# List of Figures

# List of Tables

# Acknowledgements

I would like to thank the many people who have provided support, encouragement, advice and direction during the course of my PhD.

First of all, I wish to thank my family: my parents - Gillian and Stephen - for always being there for support whenever I needed them throughout my time; my little sister - Amy - for weekly telephone conversations that put a smile on my face; and my Nan - Rita - who is also very dear to me.

I want to thank my advisor, Dr. Daniel Ueltschi, for his advice and encouragement in undertaking the work for this thesis. He is always generous with his time and gives detailed advice for improving my work in various ways, from presenting talks at conferences to advice on the presentation of papers.

I want to thank Dr. Sabine Jansen and Dr. Dimitrios Tsagkarogiannis for useful collaborations, leading towards the paper, which comprises the final chapter of this thesis. In particular Dr. Sabine Jansen provided me with valuable feedback and advice about my first paper.

There are many other academics whom I have met at various conferences or seen their courses, who have also provided me with inspiration and knowledge during those times.

I also would like to thank Prof. David Brydges and Prof. Roman Kotecký for agreeing to examine my thesis and for the helpful and constructive comments and recommendations to improve my thesis in the final editing phase.

There are many other great PhD students I have met during my time at Warwick and also at various conferences I have attended during my PhD.

I particularly wish to thank my friend and fellow Ph.D. student Kat Rock

for the immediate friendship upon my arrival at Warwick and especially for the assistance she provided with lifts when I injured myself!

I wish to also thank the various staff at the Department of Mathematics at the University of Warwick, who are always friendly and helpful.

I have also been very fortunate to visit many great places during my time as a Ph.D. student and obtaining funding for doing so.

I particularly note the external funding received to attend the following conferences:

- Probability at Warwick 2010
- Current topics in Mathematical Physics at the ESI in Vienna
- The Prague Summer Schools of both 2011 and 2013
- The week spent at the Trimester programme at the Hausdorff Institute in Bonn, which made the collaborative paper possible
- The short visit to the ETH in Zürich for talks on Mathematical Physics
- The spring school on threshold phenomena at the IHP in Paris
- The Current Topics in Mathematical Physics summer school at CIRM in Marseille
- The winter school on spatial models in statistical mechanics at Darmstadt
- The various workshops funded by the University of Warwick and EPSRC

The research conducted for this PhD was fully funded by EPSRC grant EP/G056390/1.

# Declarations

This thesis is submitted to the University of Warwick in support of my application for the degree of Doctor of Philosophy. It has been composed by myself and has not been submitted in any previous application for any degree.

The work presented was carried out by the author except in the cases outlined below:

Chapters I and II are, in general, expository. The material is not new, but it is presented to motivate the further chapters.

Chapters V and VI are adapted from the book [BLL98] and related papers mentioned in the chapters. It is a mixture of material that is already known and of connections original to this thesis towards statistical mechanics.

Chapter VII involves expository material on graph-tree identities and matroids, which is not original. It is presented so that the theory can be further extended to the setting of two-connected graphs.

Chapter IX involves a collaborative piece of work with Jansen Tsagkarogiannis and Ueltschi

Material from papers is also presented within this thesis, but a clear indication is given when a theorem is to be attributed to someone else.

Parts of this thesis have been published by the author:

There is one paper which comprises the majority of Chapter IV, which has been published [Tate13]: S. J. Tate, *Virial Expansion Bounds* Journal of Statistical Physics, 153, 2, 325–338 (2013)

There is a collaborative piece, presented as part of Chapter IX, which has been accepted and should be published soon. [JTTU14] S. Jansen, S. J. Tate,

D. Tsagkarogiannis, D. Ueltschi, *Multispecies Virial Expansions* Communications in Mathematical Physics, 330, 2, 801–817 (2014)

There is a recently submitted paper, which forms most of Chapter III [Tate14]: S. J. Tate, *A Solution to the Combinatorial Puzzle of Mayer's Virial Expansion* arXiv:1402.2119

No work presented in this thesis has been submitted for a prior examination or been published before the commencement of the PhD studies.

# Notation

## Basic Notation

| | |
|---|---|
| $\mathbb{N}$ | The set of natural numbers (exclusive of 0) |
| $\mathbb{N}_0$ | The set of natural numbers inclusive of 0 |
| $\mathbb{Z}$ | The set of integers |
| $\mathbb{R}$ | The set of real numbers |
| $\mathbb{C}$ | The set of complex numbers |
| $\emptyset$ | The empty set |
| $A + B$ | The disjoint union of the sets $A$ and $B$ |
| $[N]$ | The set $\{1, \cdots, N\}$ for $N \in \mathbb{N}$ |
| $[N]^{(2)}$ | The set of unordered pairs in $[N]$ $\{\{i, j\} \mid i \neq j \in [N]\}$ |
| $\chi_S$ and $\mathbb{1}_S$ | The indicator function on a set $S$ |
| $\mathbf{v}$ | A vector quantity |
| $\Re(z)$ | The real part of $z \in \mathbb{C}$ |
| $\Im(z)$ | The imaginary part of $z \in \mathbb{C}$ |
| $[a, b]$ | For $a \in \mathbb{N}$ and $b \in \mathbb{N}$ denotes the set $\{x \mid x \in \mathbb{N}\, a \leq x \leq b\}$ |
| $(\mathbf{x})_n$ | The sequence of vectors $(\mathbf{x}_1, \cdots, \mathbf{x}_n)$ for $n \in \mathbb{N}$ |
| $(\mathbf{x})_S$ | for $S$ a general set, this denotes the sequence of vectors with subscripts in $S$ |
| $\int \mathrm{d}^D(\mathbf{y})_n$ | The integral measure $\int \mathrm{d}^D\mathbf{y}_1 \cdots \int \mathrm{d}^D\mathbf{y}_n$ |
| $\mathrm{sign}(x)$ | The function sign : $\mathbb{R} \to \{-1, 0, 1\}$, defined by: $\mathrm{sign}(x) = \begin{cases} -1 & \text{if } x < 0 \\ 0 & \text{if } x = 0 \\ 1 & \text{if } x > 0 \end{cases}$ |
| $\lvert\Lambda\rvert = \mathrm{Vol}(\Lambda)$ | Volume of Lebesgue measurable set $\Lambda$ |
| $\mathcal{S}[K]$ | Symmetric group on the set $K$ |
| $\mathrm{Res}_{z=\zeta} f(z)$ | The residue of the complex function $f$ at $z = \zeta$ |
| $\mathrm{Res} f(z)$ | The residue of the complex function $f$ at $z = 0$ |
| $[z^n] f(z)$ | The coefficient of $z^n$ in the expansion of $f$ |
| $\mathbb{P}$ | A probability measure |
| $\mathcal{P}[S]$ | The power set of the set $S$ |

| | |
|---|---|
| $\delta_{i,j}$ | The identity matrix or Kronecker delta |
| $\mathbf{z^n}$ | The product $\prod_{i\in I}(z_i^{n_i})$, where $I$ is the index set for the vector |
| $\mathbf{n!}$ | The product $\prod_{i\in I}(n_i!)$, where $I$ is the index set for the vector |
| $[\mathbf{z^n}]f(\mathbf{z})$ | The coefficient of $\mathbf{z^n}$ in the expansion of $f(\mathbf{z})$ |
| $\lvert\mathbf{n}\rvert$ | The 1-norm of the vector of natural numbers $\mathbf{n}$, defined by $\sum_{i\in I} n_i$ |
| $[\mathbf{n}]$ | Is a label set, denoting the set $[\lvert\mathbf{n}\rvert]$ with the canonical colouring |
| $\lvert\alpha_{ij}\rvert_{i,j\in J}$ | The determinant of matrix with entries $\alpha_{ij}$ and index set $J$ |

All integrals are taken with respect of Lebesgue measure unless stated.

## Thermodynamic Functions

| | |
|---|---|
| $x_i$ | Spatial point for particle $i$ |
| $q_i$ | Generalised coordinate for particle $i$ |
| $p_i$ | Conjugate momentum for particle $i$ |
| $D$ | The dimension of the space |
| $N$ | The number of particles in finite particle system |
| $\Phi$ | The function $\Phi : \mathbb{R}^D \to \mathbb{R}$, the symmetric translation invariant pair potential in $x_i$ |
| $U$ | denotes the function $U : \mathbb{R}^D \to \mathbb{R}$, the symmetric translation invariant pair potential in $q_i$ |
| $\mathcal{U}$ | The potential energy (usually the sum of pair potentials but can be more general) |
| $H(\{q_i, p_i\})$ | The Hamiltonian of the system |
| $\Lambda$ $V$ | The finite volume of space, used as a subscript to denote finite versions of thermodynamic variables |
| $f_{i,j}$ | The Mayer $f$-function, defined by $f_{i,j} := \exp(-\beta\Phi(x_i - x_j)) - 1$ |
| $P$ | The pressure |
| $\beta = \frac{1}{kT}$ | Inverse temperature |
| $k$ | Boltzmann's constant |
| $T$ | Temperature |
| $\rho$ | Density |
| $z = e^{\beta\mu}$ | Fugacity or activity |
| $\mu$ | The chemical potential |
| $Z_N$ | Canonical Ensemble Partition Function |
| $\Xi(z)$ | Grand Canonical Ensemble Partition Function |

| | |
|---|---|
| $h$ | Parameter carrying dimensions of action, used to make partition functions non-dimensional |
| $\lambda$ | Thermal wavelength $\left(\frac{\beta h^2}{2\pi m}\right)^{\frac{D}{2}}$ |
| $v$ | The reduced volume $\frac{V}{N}$ |
| $f_\beta(\rho)$ | Thermodynamic free energy |
| $f_{\beta,\Lambda}(N)$ | Finite volume free energy |

## Combinatorial Notation

| | |
|---|---|
| $\mathcal{G}[N]$ | The collection of simple graphs (no loops or repeated edges) on vertex set $[N]$ for $N \in \mathbb{N}$ |
| $\mathcal{G}[V]$ | The collection of simple graphs (no loops or repeated edges) on vertex set $V$ |
| $g$ | A particular graph |
| $E(g)$ | The edge set of the graph $g$ |
| $e(g)$ | The cardinality of the edge set of the graph $g$ |
| $V(g)$ | The vertex set of the graph $g$ |
| $\mathcal{C}[N]$ | The collection of connected graphs on vertex set $[N]$ for $N \in \mathbb{N}$ |
| $\mathcal{C}[V]$ | The collection of connected graphs on vertex set $V$ |
| $\mathcal{B}$ | For two-connected graph versions |
| $\mathfrak{a}$ | For trees |
| $\mathcal{A}$ | For rooted trees |
| $\mathcal{E}$ or SET | The set species |
| $\mathcal{P}$ | The power set species |
| $\mathcal{S}$ and PER | The permutation species |
| $\mathcal{F}_W(z)$ | The exponential generating series for species $\mathcal{F}$ with weight $W$ |
| $B_k$ | The number of blocks of size $k$ in a graph |
| $B_{\geq k}$ | The number of blocks of size at least $k$ in a graph |
| $\mathcal{L}$ | The species of linear orderings |
| $\mathcal{F}$ | The species of forests |
| $d(i)$ | The degree of vertex $i$ |
| $d_1(i)$ | The graph distance from the vertex labelled $i$ to the vertex labelled 1 |
| $g \setminus S$ | For a graph $g = (V(g), E(g))$ and set $S$, denotes the graph with vertex set $V(g) \setminus S$ and edge set $E(g) \setminus S^{(2)}$ |
| $\mathrm{Par}[S]$ | The collection of partitions of the set $S$ |

## Polymer Notation

| | |
|---|---|
| $\gamma$ | Polymer |
| $\mathfrak{P}$ | A countably infinite collection of polymers |
| $\mathfrak{K}$ | The polymer collection for the subset gas |
| $\mathfrak{F}$ | A finite collection of polymers |
| $\Phi(\gamma)$ or $z_\gamma$ | Complex polymer activity |
| $\Psi(\gamma)$ | Positive real polymer activity |
| $\mathcal{Z}(L, \Phi)$ | Polymer partition function on $L \subset \mathfrak{P}$ and $\Phi$ the polymer activity |
| $V(\gamma_i, \gamma_j)$ | Polymer pair potential |
| $\iota$ | Incompatibility relation |

## Cluster and Virial Expansion Notation

| | |
|---|---|
| $W_N(1, \cdots, N)$ | The Mayer function, with arguments being variables with indices corresponding to the given labels |
| $U_N(1, \cdots, N)$ | The (connected) Ursell function, with arguments being the variables corresponding to the given labels |
| $V_N(1, \cdots, N)$ | The (two-connected) Husimi function, with arguments being the variables corresponding to the given labels |
| $l_n$ | The number of connected graphs on $n$ vertices |
| $w_n$ | The number of two-connected graphs on $n$ vertices |
| $\mathcal{R}_{\text{Vir}}$ | The radius of convergence of the virial expansion |
| $\mathcal{R}_{\text{May}}$ | The radius of convergence of the Mayer or cluster expansion |
| $b_l$ | The cluster coefficient for the expansion $\beta P = \sum_{l=1}^{\infty} b_l z^l$ |
| $\beta_i$ | The $(i+1)$th irreducible integral, the integral of $V_{i+1}$ |
| $c_n$ | The virial coefficient for the expansion $\beta P = \sum_{n=1}^{\infty} c_n \rho^n$ |
| $J$ | An index set for particle type in the multispecies expansions |
| $b(\mathbf{n})$ | The multispecies cluster coefficients $\beta P = \sum_{\mathbf{n} \in \mathbb{N}_0^J} b(\mathbf{n}) \mathbf{z}^{\mathbf{n}}$ |
| $c(\mathbf{n})$ | The multispecies virial coefficients $\beta P = \sum_{\mathbf{n} \in \mathbb{N}_0^J} c(\mathbf{n}) \boldsymbol{\rho}^{\mathbf{n}}$ |

In the finite volume case, we usually indicate the volume variable $\Lambda$ as a subscript.

## Matroid Notation

| | |
|---|---|
| $\mathcal{M}$ | A matroid |
| $m$ | A multiplicative function |
| $\mathcal{N} = (\mathcal{M}, m)$ | An arithmetic matroid |
| $E$ or $E(\mathcal{M})$ | The ground set of a matroid |
| rk | The rank function |
| $\mathcal{I}(\mathcal{M})$ | The collection of independent sets |

# Abstract

The interpretation of cluster and virial expansions as weighted exponential generating functions of connected and two-connected graphs respectively was given by Mayer in the 1940s. Combinatorial approaches, either through the tree graph identities, introduced by Brydges, Battle and Federbush, or the fixed point equations of Kotecký-Preiss, generalised by Fernández-Procacci, have led to results pertaining to the convergence of these series, notably in the case of the cluster expansion. Recent interest in these expansions has been stimulated by the connection to Joyal's combinatorial species of structure, presented both in the work of Leroux and his collaborators and Faris. Virial expansions have also gained renewed interest through the Canonical Ensemble methods of Pulvirenti and Tsagkarogiannis, through which the convergence conditions of Lebowitz and Penrose are obtained.

This thesis obtains combinatorial interpretations of the cancellations in the virial expansions for the one-particle hardcore gas and the Tonks gas. Improved bounds are obtained for the virial expansion, through an original approach, depending on cluster coefficient bounds. Separate bounds are also found by using the cluster coefficient bounds of Poghosyan and Ueltschi. Furthermore, a generalisation to multispecies expansions, through Lagrange-Good inversion, is given, providing Kotecký-Preiss type conditions of convergence for the multispecies virial expansion.

The tree-graph expansions are analysed in the context of these results and are used to understand the key structure necessary to replicate such bounds for virial expansions.

# Introduction

A central question of statistical mechanics is the calculation of the thermodynamic properties of a given fluid starting from the forces between the molecules, that is obtaining a macroscopic description of a system from the microscopic description of particles. One wishes to obtain the macroscopic description from a quantum mechanical or classical description of the system, which is dependent on the intermolecular forces. Two tools useful in doing so at high temperature and low density regimes are the cluster and virial expansions.

Virial and cluster expansions arise from considering the definition of pressure in the Grand Canonical ensemble. Pressure may be written in terms of an activity coefficient $z = e^{\beta\mu}$, where $\beta$ is the inverse temperature and $\mu$ the chemical potential. This is expanded as a power series, which is known as the cluster expansion. Alternatively, we may take density $\rho$ as the parameter for the system and obtain an expansion for pressure in terms of density, which is known as the virial expansion. The two expansions indicate the relationship between pressure and two macroscopic variables of the system: the chemical potential $\mu$; and the density $\rho$, for which the coefficients are understood in terms of the microscopic nature via the pair potential. This thesis focuses upon obtaining a clearer understanding of cluster and virial expansions, both combinatorially and in terms of improved bounds.

The original contribution provided by this thesis comprises of three main parts. The first is an improved bound on virial coefficients and the radius of convergence of the virial expansion following from the development of a general approach to using cluster expansion bounds in the virial expansion. New bounds are also obtained from the alternative cluster expansion bounds achieved by Poghosyan and Ueltschi [PoUe09], through the modification made by Procacci [Pro07] to the tree graph expansions of Brydges and Federbush [BrFe78]. The two models of the one-particle hardcore gas and the Tonks gas are described and a combinatorial interpretation provided for the cancellation of the coefficients in the corresponding virial expansions, answering the challenge posed in the paper of Ducharme Labelle

and Leroux [DLL07]. It is anticipated that this understanding should lead to improved bounds. The final key original contribution is the extension of the cluster and virial expansions to the multispecies case. The originality of the extension is in the approach using combinatorial species and Lagrange-Good inversion with Kotecký-Preiss style techniques to obtain bounds on virial coefficients.

Chapter I introduces the concept of cluster expansions, specifically the Mayer expansion and how the coefficients of these expansions are understood as weighted connected graphs. The second part of the chapter focuses upon different polymer models as instructive examples. These are the subset gas of Gruber and Kunz [GK71] and the abstract polymer model introduced by Kotecký and Preiss [KoPr86]. The chapter explains the collection of fixed point equations, which provides criteria for the convergence of the cluster expansion for abstract polymer models. This is developed from those given by Kotecký Preiss [KoPr86] to Dobrushin [Dob96] and finally into the paper of Fernández and Procacci [FePr07]. The separate direction of Gruber and Kunz following Kirkwood Salsburg equations [KiSa53] has been unified with this fixed point approach and this is explained towards the end of the chapter. Furthermore, explicit coefficient bounds developed in the work of Groeneveld and found in [McCoy10] are also conveyed, indicating the tight result for positive potentials. Further approaches to cluster expansion bounds are indicated in later chapters.

Chapter II introduces the virial expansion and interprets the approximation provided by the van der Waals Maxwell theory of the equation of state. The chapter gives Mayer's proof that the virial coefficients are weighted two-connected graphs, as found in [McCoy10] highlighting the importance of the irreducible integrals. Using the Kirkwood Salsburg equation and the algebraic approach of Ruelle [Rue69], a simple bound is presented. The connections to the cluster expansion are made explicit here. The final topic is the Canonical Ensemble calculations made by Pulvirenti and Tsagkarogiannis [PT12], which achieve bounds on the virial expansion and re-derive the weighted two-connected graph identity for the virial coefficients.

Chapter III presents two simple statistical mechanical examples: the one-particle hardcore gas; and the Tonks gas or one dimensional continuum hardcore gas. The cluster and virial expansions of these two models can be easily derived without the machinery of combinatorics. The key point is to compare what is achieved through simple derivations with what is achieved through Mayer's theorems [MMay40] relating the coefficients to weighted graphs. This leads us to the four

combinatorial identities:

$$\sum_{g\in\mathcal{C}[n]}(-1)^{e(g)}=(-1)^{n-1}(n-2)! \tag{0.0.1}$$

$$\sum_{g\in\mathcal{B}[n]}(-1)^{e(g)}=-(n-2)! \tag{0.0.2}$$

$$\sum_{g\in\mathcal{C}[n]}(-1)^{e(g)}\operatorname{Vol}(\Pi_g)=(-1)^{n-1}n^{n-1} \tag{0.0.3}$$

$$\sum_{g\in\mathcal{B}[n]}(-1)^{e(g)}\operatorname{Vol}(\Pi_g)=-n(n-2)! \tag{0.0.4}$$

where $\Pi_g$ is the polytope defined by:

$$\Pi_g:=\{(\mathbf{x})_{[2,n]}\in\mathbb{R}^{n-1}|\ |x_i-x_j|<1\,\forall\{i,j\}\in E(g)\} \tag{0.0.5}$$

where $x_1=0$.

The identities have been explained in the connected graph case by Bernardi [Ber08] and these identities appear in the paper of Ducharme Labelle and Leroux [DLL07]. This chapter explains the connected graph case and then proceeds to explain the case for two-connected graphs found in [Tate14]. It is anticipated that this understanding will prove useful in developing general bounds for weighted two-connected graph sums. Another key point is that, for the cluster expansion with repulsive interactions, the Tonks gas and the one-particle hardcore gas represent two extremes on the values of the cluster coefficients.

Chapter IV gives an overview of current abstract bounds for the coefficients and radius of convergence of the virial expansion. It involves key bounds from Groeneveld for positive potentials and an improvement and generalisation of the bounds provided by Lebowitz and Penrose [LePen64], which appear in the paper [Tate13]. The bounds of Morais and Procacci [MoPr13] are also indicated and the links between the various bounds are made explicit. Comparisons are made between the bounds and a general derivation of virial expansion bounds from cluster coefficient bounds is given. As an application, the bounds appearing in the paper of Poghosyan and Ueltschi [PoUe09], which are inspired by the work of Procacci [Pro07], are used in this general derivation to obtain a new alternative formulation of a virial coefficient bound. This chapter ends by looking forward to directions that can be taken to improve our understanding of the virial expansion, which are described in later chapters.

Chapter V follows the work of Faris [Far10,Far11], Joyal [Joy81] and Bergeron

Labelle and Leroux [BLL98]. It provides an introduction to the theory of combinatorial species of structure, setting the context of it as a functor from the category of finite sets with bijections to itself. In this chapter, the development of the subject is motivated through the application to cluster and virial expansions. This theory is further developed in Chapter VI, which presents more advanced topics specifically important for statistical mechanics. The key topics are the dissymmetry theorem and Lagrange inversion, which are important for the tree-like relationships, present between connected and two-connected graphs.

Chapter VII indicates the graph-tree identities first applied and understood in the papers by Brydges Battle Federbush and Kennedy [BaFe84, Bat84, BrFe78, Bry84, BrKe87]. This is in parallel to the Penrose construction [Pen67], which allows for the weighted sum over connected graphs to be expressed in terms of trees with a modified weight in order to effectively deal with the cancellations. The chapter concludes by considering the final generalisation to matroids by Faris [Far12a]. This is also shown in the context of the q-state Potts model by Sokal [Sok05]. The key notions of internally and externally active edges, which generalise the Penrose construction, are anticipated to provide more efficient methods of dealing with the virial expansion coefficients. The key idea is that the combinatorics of certain expressions can be simplified by reducing the objects considered to a smaller set, but where a price is paid in terms of the weights in an analytic sense. It indicates a setting for Chapter III in how it may generalise the involution relying on what may be called active edges. The understanding of the internally active edges is not present yet.

Chapter VIII develops the overall idea of Groeneveld [Gro67] that we are able to make such estimates for different types of graphs, such as through the notion of minimal two-connected graphs. It is explained that this provides difficulties in enumeration and actually finding a partition analogous to that of Penrose. The Ree-Hoover expansion is explained in the context of Chapter III.

Chapter IX gives the generalisation found in the collaborative paper with Jansen Tsagkarogiannis and Ueltschi [JTTU14] to multispecies cluster and virial expansions. The history of this expansion, through the work of Fuchs [Fuc42] and Mayer [May39, HaMa38] is explained, as are the difficulties of the generalisation of the van der Waals Equation of State [LeRo64]. The useful combinatorial tool for this is the Lagrange-Good inversion [Good60, Good65], which generalises the Lagrange and the Lagrange Bürmann inversion formulæ, which are presented in Chapter VI. A discussion of the approaches to proving this and its wider applications and connections is indicated. We obtain bounds on the fugacity parameters, which allow for the inversion of the multi-species cluster expansion into the multi-species

virial expansion. The bounds are then shown to hold in the case of a rigid polymer system. Problems arise in the case when the polymers are not assumed to be rigid and this is discussed. The application follows the paper of [PoUe09]. The context of the multispecies dissymmetry theorem for coloured objects is also explained to allow for an interpretation of the virial coefficients in particular cases. Furthermore, an analytical context of this inversion is indicated as extending an inverse function theorem for Fréchet differentiable functions in many variables.

# Chapter I

# Cluster Expansions

This chapter presents the background to the theory of cluster expansions and the historical development of the subject. The term, cluster, emphasises the fact that the coefficients depend upon weighted connected graphs. The imagery expected to be aroused by the term cluster is the notion of groups of particles close in space that only interact in small clusters. There are notable links to combinatorics and probability theory. In particular aspects of the structure of this theory can be appreciated from the point of view of combinatorial species of structure, as presented in Chapter V.

The cluster expansion was realised in the 1940's by Mayer in [MMay40]. Mayer provided a method of perturbing the grand canonical partition function as an expansion in the fugacity parameter through the use of Mayer graphs. Although the grand canonical partition function is analytic for the cases studied in this thesis, most thermodynamic functions rely on taking the logarithm of this, which causes issues of analyticity at the zeroes of the partition function. This interpretation is found in [Sok01, FePr08] among other papers. These expansions are possible for high-temperature approximations or low density approximations. The interactions, which appear within the coefficients of the expansion are also assumed to obey certain properties, which often means making them small.

Mayer first used the clever idea to transform the product of the exponential of pairwise interactions into the sum of graphs with edge weights

$$f_{i,j} = \exp\left(-\beta U(q_i - q_j)\right) - 1 \tag{I.0.1}$$

The work of Mayer establishes, in what is known as Mayer's first theorem, the connection of the coefficients of the fugacity expansion for pressure and weighted connected graphs. Mayer's method can be found in [MMay40] and a review of this is found in the book by McCoy [McCoy10].

The importance of cluster expansions within statistical mechanics, relies on the fact that we obtain analytic expansions of pressure and density, providing the fugacity, $z$, lies within the radius of convergence of the cluster expansion. Finding the radius of convergence, gives bounds on regions where the thermodynamic variables cannot undergo a phase transition, if we understand phase transitions as corresponding to points of non-analyticity of the thermodynamic variables.

This approach, however, was not developed much further until the 1960s, when Groeneveld in [Gro62] produced further results on the radius of convergence. Penrose achieved independent bounds on the radius of convergence of cluster expansions, through the Kirkwood Salsburg equations in [Pen63a]. Ruelle developed an algebraic approach to cluster expansions complementing the use and interpretation of Kirkwood Salsburg equations and the Mayer and Montroll equations in his book, [Rue69]. The main ideas on this topic are found in the earlier paper of Ruelle [Rue64].

The subset gas model as a polymer system was used by Gruber and Kunz in the paper [GK71] to obtain estimates on the radius of convergence of cluster and virial expansions. Kotecký and Preiss then presented a method of understanding the convergence of cluster expansions for (abstract) polymer models without needing to resort to using Kirkwood Salsburg equations or tree-estimation methods in [KoPr86]. Kotecký Preiss-like conditions have been applied in further work concerning the convergence of cluster expansions to understand it applied to various different models. The most notable paper with a comprehensive interpretation of how Kotecký Preiss like conditions may be applied to different models in statistical mechanics is the paper by Poghosyan and Ueltschi [PoUe09]. Furthermore, the results of this paper were instructive in the paper on multispecies expansions [JTTU14], which is discussed in Chapter IX. Dobrushin in [Dob96] and Fernández and Procacci in [FePr07] detail how to weaken this condition, in order to obtain better bounds, but with the same style as Kotecký Preiss, by presenting the bounds as a fixed point equation.

The inductive approach by Dobrushin is further refined in the paper by Miracle-Solé [Mir00] and has been unified with the approach taken by Gruber and Kunz in using Kirkwood Salsburg equations in the paper by Bissacot Fernández and Procacci [BFP10]. They indicate how the approach obtains the improved Fernández-Procacci bound found in [FePr07].

The bounds given in the specific case of positive (repulsive) potentials by Groeneveld [Gro62] are the best possible without further information. This is highlighted by two simple, yet instructively important examples: that of the single particle hard core gas and the Tonks gas, as are presented in Chapter III. If we want

better bounds than those achieved by Groeneveld, then we would need to make further assumptions about the potential. The actual expressions give also interesting combinatorial identities, when we make the link to weighted connected graphs. Bernardi, in [Ber08], gives combinatorial proofs of the two identities we obtain from these models. It is anticipated that these will help understand the cancellations in more general models better.

Furthermore, cluster expansions have been applied to other contexts outside of statistical mechanics. The methods of cluster expansions are applicable in probability, through resolvent expansions [AbRi94]. They also make an important appearance in conformal quantum field theory and the loop vertex expansion of Rivasseau and Wang [RiWa10a, RiWa10b]. Many of the key ideas of such expansions are found and used in various aspects of quantum field theory in attempting to make sense of renormalisation. Furthermore, an extension of the method by Fernández and Procacci is presented in the paper by Temmel [Tem12]. Cluster expansions are also relevant for chromatic polynomials [FePr08] and the Lovasz local lemma [BFPS11]. In this chapter, however, the important techniques and concepts of cluster expansions are presented with the key purpose of improving the understanding and bounds used for the virial expansion.

## I.1 The Partition Function

This section presents background material to statistical mechanics. The material is not novel and is found in references [Fey72, Hill56, LiPi80, LLP84, MMay40, McCoy10]. The presentation is made so that it leads naturally towards the notion of cluster expansions in the subsequent sections.

In statistical mechanics, a fundamental object is the Hamiltonian $H(\{p_i, q_i\})$, which depends on the collection of generalised coordinates $q_i$ and conjugate momenta $p_i$ of the particles. The partition function arises from summing/integrating over the whole phase space of the system, with an appropriate weight. This gives the relative probabilities of finding a system in a given state. The summation over all possible states with a fixed number, $N$, of particles is called the *canonical partition function*, written $Z_N$. For a $D$-dimensional system, the partition function is:

$$Z_N = \frac{1}{N! h^{DN}} \int_{V^N} \int_{\mathbb{R}^N} \exp(-\beta H_N(\{p_i, q_i\}) \mathrm{d}^N \mathbf{p} \mathrm{d}^N \mathbf{q} \tag{I.1.1}$$

The $h$ factor is necessary to make the quantity dimensionless. It carries the dimension of the action.

The Hamiltonian can usually be decomposed into the ideal part, involving only the momenta, and the interacting part:

$$H_N(\{p_i, q_i\}) = \sum_{i=1}^{N} \frac{p_i^2}{2m} + \mathcal{U}(\{q_i\}) \tag{I.1.2}$$

The partition function gives the total weight of the system. If we divide the weight of a point in phase space by the partition function we obtain a probability density.

If we drop the assumption of a fixed number of particles and control particle number through the chemical potential $\mu$, we have the grand canonical partition function. This is formed by taking the sum over all $N$-particle canonical partition functions, weighted by $z^N$. This free movement between the two ensembles is justified in certain cases by the notion of 'equivalence of ensembles'.

$$\Xi(z) = \sum_{N=0}^{\infty} Z_N z^N \tag{I.1.3}$$

The chemical potential in the Grand Canonical Ensemble gives a control over particle number and density, making states with a larger number of particles less or more likely, depending on the value of $\mu$. We can give each quantity in the total phase space of an unrestrained number of particles the weight in (I.1.1) corresponding to fixed $p_i$, $q_i$ and $N$ and divide by the grand canonical partition function, (I.1.3), in order to get a probability distribution. This is known as the *Gibbs Distribution.* The Gibbs distribution is a useful starting point for defining thermodynamic quantities as expectations with respect to this measure. For example, we can compute the expected number of particles:

$$\begin{aligned} \langle N \rangle &= \frac{\sum_{N=0}^{\infty} N Z_N z^N}{\Xi(z)} = \frac{z \frac{\partial}{\partial z} \Xi(z)}{\Xi(z)} \\ &= z \frac{\partial}{\partial z} \ln \Xi(z) = kT \left( \frac{\partial \ln \Xi(z)}{\partial \mu} \right)_\beta \end{aligned} \tag{I.1.4}$$

### I.1.1 The Ideal Gas

In the ideal gas case, where the interaction potential $\mathcal{U}(\{q_i\})$ in (I.1.2) is set to zero, we have

$$H_N(\{p_i, q_i\}) = \sum_{i=1}^{N} \frac{p_i^2}{2m} \tag{I.1.5}$$

For the Canonical Ensemble, we integrate the momenta as Gaussian variables and realise that the spatial integrals give the volume of the space, to obtain:

$$Z_N^{\text{ideal}} = \frac{1}{N!}\left(\frac{2m\pi}{\beta h^2}\right)^{\frac{DN}{2}} V^N \tag{I.1.6}$$

Defining the thermal wavelength $\lambda$ by:

$$\lambda = \left(\frac{\beta h^2}{2m\pi}\right)^{\frac{D}{2}} \tag{I.1.7}$$

We can also calculate the grand canonical partition function:

$$\Xi(z) = \sum_{N=0}^{\infty} \frac{1}{N!}\frac{z^N}{\lambda^N} V^N = \exp\left(\frac{z}{\lambda}V\right) \tag{I.1.8}$$

The thermodynamic pressure is defined by: $\beta P = \lim\limits_{V\to\infty} \dfrac{\ln(\Xi)}{V}$. In this case:

$$\beta P = \frac{z}{\lambda} \tag{I.1.9}$$

In the Canonical Ensemble we have the definition for pressure:

$$P = -\left(\frac{\partial A}{\partial V}\right)_T \tag{I.1.10}$$

where $A$ is the *Helmholtz free energy*, defined by:

$$A = -\frac{1}{\beta}\ln Z_N \tag{I.1.11}$$

Hence we have:

$$\ln Z_N = N \ln\left(\left(\frac{2m\pi}{\beta h^2}\right)^{\frac{D}{2}} V\right) \tag{I.1.12}$$

From (I.1.10) and (I.1.11), we have the expression for pressure:

$$\begin{aligned} \beta P &= \frac{\partial}{\partial V}\ln Z_N \\ &= \frac{\partial}{\partial V}\left(N \ln\left(\left(\frac{2m\pi}{\beta h^2}\right)^{\frac{D}{2}} V\right)\right) \\ &= \frac{N}{V} \end{aligned} \tag{I.1.13}$$

Letting $\rho = \frac{N}{V}$, the density, we have: $\beta P = \rho$. Alternatively, we have the ideal gas

law:

$$VP = NkT \tag{I.1.14}$$

## I.2 Fugacity Expansions of the Pressure - Mayer's First Theorem

Mayer's first theorem is found in [MMay40]. This section gives an alternative derivation of the theorem.

The ideal gas case is not completely realistic for most gases seen in nature. In order to create a more realistic model, we include the interaction potential $\mathcal{U}(\{q_i\})$

**Assumption 1** (Assumptions on the Interactions)**.** *We assume that interactions are only pairwise, symmetric in the two variables and translation and rotation invariant. The interaction can then be written in the form:*

$$\mathcal{U}(\{q_i\}) = \sum_{i<j} U(q_i - q_j) \tag{I.2.1}$$

This assumption may be justified by the fact that in the gas phase, it is unlikely that many particles will all be close to each other to interact, since the density is sufficiently small. Pairwise interactions will therefore be rare and interactions involving many more particles will be even more rare. These occur so infrequently that we may assume them to be negligible for this first approximation of a model of the gas. We thus justify focusing on only pairwise interactions, by assuming many body interactions to be negligible. We can now write the grand canonical partition function as:

$$\Xi(z) = \sum_{N=0}^{\infty} \frac{z^N}{N!\lambda^N} \int_V \mathrm{d}q_1 \cdots \int_V \mathrm{d}q_N \, \exp\left(-\beta \sum_{1\le i<j\le N} U(q_i - q_j)\right) \tag{I.2.2}$$

From the definition $\beta P = \lim_{V\to\infty} \dfrac{\ln(\Xi)}{V}$ and (I.2.2), we realise we need to understand the logarithm of a rather complicated function. However, Mayer [MMay40] introduced a useful transformation, which allows us to understand the function

$$\exp\left(-\beta \sum_{1\le i<j\le N} U(q_i - q_j)\right)$$

as a sum over graphs, by rewriting $\exp(-\beta U(q_i - q_j)) = 1 + f_{i,j}$. We recognise that:

$$\exp\left(-\beta \sum_{1\le i<j\le N} U(q_i - q_j)\right) = \prod_{1\le i<j\le N} \exp(-\beta U(q_i - q_j)) = \prod_{1\le i<j\le N} (1 + f_{i,j}) \tag{I.2.3}$$

We expand the RHS of (I.2.3) to get:

$$\prod_{1\le i<j\le N} (1 + f_{i,j}) = 1 + \sum_{1\le i<j\le N} f_{i,j} + \sum\sum f_{i,j} f_{k,l} + \cdots \tag{I.2.4}$$

More precisely we can see that for each unordered pair $\{i, j\}$ it either appears in one of the terms on the RHS of (I.2.4) or it doesn't (i.e. we choose 1 from the appropriate factor). It must be that this is the sum over all possible choices of subsets of $[n]^{(2)}$. The unordered pairs are interpreted as (undirected) edges in a graph with vertex set $[n]$. Hence our sum on the RHS can be viewed as the sum over all possible graphs with any number of edges. We have the interpretation:

$$\prod_{1\le i<j\le N} (1 + f_{i,j}) = \sum_{G\in\mathcal{G}[N]} \prod_{\{i,j\}\in E(G)} f_{i,j} \tag{I.2.5}$$

where the product over the empty set is 1.

The fugacity expansion for the grand canonical partition function is thus:

$$\Xi(z) = \sum_{N=0}^{\infty} \frac{z^N}{N!\lambda^N} \int_V \mathrm{d}q_1 \cdots \int_V \mathrm{d}q_N \sum_{g\in\mathcal{G}[N]} \prod_{\{i,j\}\in E(g)} f_{i,j} \tag{I.2.6}$$

We wish to express (I.2.6) as an exponential, so that the logarithm of (I.2.6) can be easily recognised.

**Definition** (Connected Graphs and Connected Components)**.** *A connected subgraph $c = (V', E')$ of a graph $g$ has vertex set $V' \subset V(g)$ and an edge set $E' \subset E(g)$, which consist of all edges in $E(g)$ such that both endpoints are in $V'$. It also has the property that you can go from any vertex to another by following a path along the edges included in the subgraph. A connected component is a maximal such subgraph.*

The formulation of Mayer's First Theorem is:

**Theorem I.2.1** (Mayer's First Theorem)**.** *For a model satisfying the above assumptions, one may write the (finite volume) pressure function as:*

$$\beta P_V = \frac{1}{V} \sum_{N=1}^{\infty} \frac{z^N}{N!\lambda^N} \sum_{g\in\mathcal{C}[N]} \prod_{i\in V(g)} \left(\int_V \mathrm{d}q_i\right) \prod_{\{i,j\}\in E(g)} f_{i,j} \tag{I.2.7}$$

**Discursive Derivation of Mayer's First Theorem**

We realise that for any graph $G \in \mathcal{G}[n]$, we can partition the set $[n]$ into subsets of vertices representing connected components. This provides a decomposition of the graph into connected components. We recognise that for our graph functions, the functions 'multiply' over the edges in the graph and thus over connected components or 'clusters'.

Considering a term involving connected components $C_1$ and $C_2$, there are no $f_{i,j}$ terms in the corresponding product, where $i \in C_1$ and $j \in C_2$ and vice versa as the two components are disconnected. Thus we realise that this graph function is just the product of the two graph functions corresponding to components $C_1$ and $C_2$. We can thus write the sum over $\mathcal{G}[N]$ as first a sum over partitions $\Gamma = \{\gamma_1, \cdots, \gamma_k\}$ of $[N]$, which represent the connected components of the graph, and then over products of the sums of connected graphs in each $\mathcal{C}[\gamma_i]$.

We define the graph weight $w$:

$$w(g) = \prod_{\{i,j\} \in E(g)} f_{i,j} \tag{I.2.8}$$

So that:

$$\sum_{g \in \mathcal{G}[n]} w(g) = \sum_{\substack{\Gamma = \{\gamma_1, \cdots, \gamma_k\} \\ \Gamma \in \mathrm{Par}[n]}} \prod_{i=1}^{k} \sum_{c \in \mathcal{C}[\gamma_i]} w(c) \tag{I.2.9}$$

We define the integrated graph weight $W$ as:

$$W(g) := \prod_{i \in V(g)} \left( \int_V dq_i \right) w(g) \tag{I.2.10}$$

We have the same relations (I.2.9) for $W(g)$ as for $w(g)$, since we just need to interchange finite sums with intergrals and realise that the integrals factor for the product.

The graph functions $W(g)$ are independent of the particular labels given and so we can use just the size of the sets in the partition.

The number of partitions of a set of size $N$ into sets of size $\{N_i\}_{i=1}^k$ is the multinomial coefficient: $\binom{N}{N_1 \cdots N_k} = \frac{N!}{\prod_{i=1}^{k} (N_i!)}$. Hence, we can sum over $\{N_i\}_{i=1}^k$ satisfying $\sum_{i-1}^{k} N_i = N$, providing we include the multinomial coefficients.

We also note that it is possible to write the sets $\{N_i\}_{i=1}^k$ as a sequence, $(N_i)_{i=1}^k$ and remember to divide by $k!$ to get the appropriate correspondences, since

sequences can be permuted.

We obtain:

$$\frac{1}{N!}\sum_{g\in\mathcal{G}[N]} W(g) = \sum_{k=1}^{N}\frac{1}{k!}\sum_{\substack{(N_i)_{i=1}^{k}\\ \sum_{i=1}^{k} N_i = N}}\prod_{i=1}^{k}\left(\frac{1}{N_i!}\sum_{c\in\mathcal{C}[N_i]} W(c)\right) \tag{I.2.11}$$

Summing both sides of (I.2.11) over $N$, from 0 upwards, with a factor $\frac{z^N}{\lambda^N}$ for each $N$, we obtain the expression:

$$\Xi(z) = 1 + \sum_{N=1}^{\infty}\sum_{k=1}^{N}\frac{1}{k!}\sum_{\substack{(N_i)_{i=1}^{k}\\ \sum_{i=1}^{k} N_i = N}}\prod_{i=1}^{k}\left(\frac{z^{N_i}}{N_i!\lambda^{N_i}}\sum_{c\in\mathcal{C}[N_i]} W(c)\right) \tag{I.2.12}$$

We notice that we may interchange the sums over $N$ and $k$ as:

$$\sum_{N=1}^{\infty}\sum_{k=1}^{N} = \sum_{k=1}^{\infty}\sum_{N=k}^{\infty} \tag{I.2.13}$$

and realise that the sum over sequences is just a product of sums with the restraint, $\sum_{i=1}^{k} N_i = N$, which can be removed upon taking the sum over $N$, giving infinite upper limits to the sums on the $N_i$.

This gives us:

$$\Xi(z) = 1 + \sum_{k=1}^{\infty}\frac{1}{k!}\prod_{i=1}^{k}\left(\sum_{N_i=1}^{\infty}\frac{z^{N_i}}{N_i!\lambda^{N_i}}\sum_{c\in\mathcal{C}[N_i]} W(c)\right) \tag{I.2.14}$$

Defining the generating series:

$$A(z) = \sum_{N=1}^{\infty}\frac{z^N}{N!\lambda^N}\sum_{c\in\mathcal{C}[N]} W(c) \tag{I.2.15}$$

allows us to rewrite (I.2.14) as:

$$\Xi(z) = 1 + \sum_{k=1}^{\infty}\frac{1}{k!}(A(z))^k = \exp(A(z)) \tag{I.2.16}$$

This gives us an expression for the (finite-volume) pressure:

$$\beta P_V = \frac{A(z)}{V} \tag{I.2.17}$$

and so we have that it is the sum over weighted connected graphs or clusters, giving a cluster expansion.

**Remark 1** (Volume Dependence)**.** *We note here that in the infinite volume limit the function $A(z)$ is proportional to the volume $V$. This is due to the weight function $W(g)$ being proportional to the volume, when we take the thermodynamic limit. We see from (I.2.10) that we have integrals over each coordinate in volume. The interactions contained in $w(g)$ are all symmetric pair potentials depending only on the relative distances between the two coordinates. This means one of the coordinates is redundant and so the respective integral contributes a volume factor. In the finite volume case, this is not true due to the effect at the boundary of the region. If we consider the integrand $w(g)$ and we integrate all of the variables excepting one, in finite volume, this will depend on the location of the final coordinate that hasn't been integrated over. The pair potentials depend only on the distance from this point and so the location is important when one needs to consider the fact that the other particles cannot be outside of the box. This results in boundary effects and can be rectified by considering periodic boundary conditions, where the box is understood as a torus. These effects disappear in the infinite volume limit, giving the proportionality to volume.*

The theory of combinatorial species of structure (Chapter V )and the Möbius inversion formulæ provide further machinery to understand this relationship and the generalisations of the concept of connected components. The interpretation of connected components as combinatorial species can be found in Section VI.1. There is also a neat derivation of Mayer's First Theorem in [Kot06].

**Remark 2.** *In Mayer's original proof of the cluster expansion, a functional relation between simple (loopless with no multiple edges) graphs and connected graphs is obtained, which can be written as a differential equation. This identity is related to a combinatorial identity, found in (V.4.5). It is that a (simple) graph (with the 'special' or ghost vertex $N+1$) may be viewed as the partition of the vertices into the connected component containing this special point and any possible (simple) graph on the rest of the vertices. The differential equation here may be interpreted as functional equations in combinatorics. The equation is: $\mathcal{G}' = \mathcal{C}' \star \mathcal{G}$.*

## I.3 Polymer Models I - The Subset Gas à la Gruber Kunz

In [GK71], Gruber and Kunz give a physical polymer model, which has been termed the subset gas in [BFP10]. This is to emphasise this as a distinct model from the polymer model used in successive sections. The content of this section is adapted from [GK71].

The motivation, first of all, to study these physical polymer models in particular, is linked to the fact that any general classical lattice system can be reduced to a system of polymers on the same lattice. This is called the 'Association Problem', which can be found in [Rue69]. The physical polymer model is also easy to introduce and offers an insightful introduction to the abstract polymer models of the next section. Furthermore, Bissacot, Fernández and Procacci [BFP10], indicate a precise connection between the approach of Gruber and Kunz and that of Dobrushin.

We consider the polymer models in the context of the cluster expansion. In particular, we study different conditions for cluster expansion convergence and compare these, indicating possible ways forward to achieving tighter conditions and a better understanding of their nature.

The approach of Gruber and Kunz, involving Kirkwood Salsburg equations, was not widely used for some time. Most practitioners of cluster expansions focused firstly on showing that cancellations in the cluster expansion yield a majorising expansion in terms of tree diagrams. The bounds were then achieved by inductively summing over leaves in the tree expansion. This approach is explored in Chapter VII. The main component of the work of Gruber and Kunz is the algebraic approach, which already appeared in some of the earlier work of Ruelle [Rue63, Rue64]. The papers of Fernández and Procacci [FePr07] re-invoked interest in this work due to the improved bounds they achieved, which improved those of Kotecký and Preiss and Dobrushin and bore similarity to those of Gruber and Kunz. They have also been adapted into applications in the work by Poghosyan and Ueltschi [PoUe09] and to an interacting case by Procacci [Pro07].

### I.3.1 The Set Up of the Model

We have a finite set $\Lambda$ consisting of $N(\Lambda)$ 'sites'. Usually $\Lambda$ is a subset of $\mathbb{Z}^\nu$ or an appropriate lattice, exhibiting the required symmetries of the model. Polymers $\gamma$ correspond to subsets of $\Lambda$. A polymer of size $n$ is interpreted as a collection of $n$

particles placed at sites $x_i \in \gamma$ for $i \in [n]$. A collection of polymers must satisfy:

$$\Lambda = \cup_{i=1}^{k} \gamma_i \text{ with } \gamma_i \neq \emptyset \text{ and } \gamma_i \cap \gamma_j = \emptyset \text{ if } i \neq j \tag{I.3.1}$$

The state of the system is defined by a probability measure $\mathbb{P}_\Lambda(\{\gamma_1, \cdots, \gamma_n\})$ on the configuration space. This measure is characterised by the positive bounded function $\Psi(\gamma)$ defined on subsets $\gamma \subset \Lambda$ and it is interpreted as the activity of polymer $\gamma$. The (finite) collection of polymers is denoted $\mathfrak{K} = \{L \subset \Lambda | L \text{ finite } \}$. Analogously to the general definitions, we have:

$$\mathbb{P}_\Lambda(\{\gamma_1, \cdots, \gamma_k\}) = \Xi_\Lambda^{-1} \prod_{i=1}^{k} \Psi(\gamma_i) \tag{I.3.2}$$

where:

$$\Xi_\Lambda = \sum_{(\gamma_i) \in \mathrm{Par}[\Lambda]} \prod_i \Psi(\gamma_i) \tag{I.3.3}$$

We can then define pressure by:

$$\beta P_\Lambda = \frac{1}{N(\Lambda)} \ln \Xi_\Lambda \tag{I.3.4}$$

We also define correlation functions:

$$\rho_\Lambda(\gamma_1, \cdots, \gamma_p) = \begin{cases} 0 \quad \text{if } \gamma_i \cap \gamma_j \neq \emptyset \text{ for some } i \neq j \\ \Xi_\Lambda^{-1} \prod_{i=1}^{p} \Psi(\gamma_i) \sum_{(\gamma_\alpha) \in \mathrm{Par}[\Lambda \setminus \cup \gamma_i]} \prod \Psi(\gamma_\alpha) \text{ otherwise} \end{cases} \tag{I.3.5}$$

The basic properties of the correlation functions are:

- $$0 \leq \rho_\Lambda(\gamma_1, \cdots, \gamma_p) \leq 1 \tag{I.3.6}$$

- $$\rho_\Lambda(X) = \Psi(X) \frac{\partial}{\partial \Psi(X)} \ln P_\Lambda = \frac{1}{\beta} \frac{\partial}{\partial \mu(X)} \ln \Xi_\Lambda \tag{I.3.7}$$

where $\beta\mu(X) = \ln \Psi(X)$. The $\mu(x)$ can be interpreted as a chemical potential for the set $X$ as in Section I.1

### I.3.2 The Algebraic Approach to the Kirkwood Salsburg Equations

The algebraic approach of Ruelle [Rue69] is used to understand the development of Kirkwood Salsburg equations in the context of correlation functions. This allows us

to obtain bounds on the radius of convergence of the corresponding expansion. This can also be found in earlier papers of Ruelle, such as [Rue63] and [Rue64]. This is also used by Poghosyan and Ueltschi in the paper [PoUe09], where it is developed in conjunction with tree expansion bounds. The algebraic approach is also interpreted and generalised in the context of combinatorial species of structure in Section VI.7. There are other integral equations that may be found in the literature such as the Mayer Montroll equations [MMon41, May47], which are not covered here. The set up is as follows:

We have a complex vector space, $V_\Lambda$, of bounded complex functions $\Phi(X)$ defined on subsets $X \subset \Lambda$. We may view this as a commutative algebra with unit element $\mathbb{1}$ for the $\star$-product, defined by:

$$\forall \Phi_1, \Phi_2 \in V_\Lambda \;\; (\Phi_1 \star \Phi_2)(X) = \sum_{Y \subset X} \Phi_1(Y)\Phi_2(X \setminus Y) \tag{I.3.8}$$

with the unit element defined by:

$$\mathbb{1}(X) := \begin{cases} 1 & \text{if } X = \emptyset \\ 0 & \text{if } X \neq \emptyset \end{cases} \tag{I.3.9}$$

**Remark 3.** *The formulation of the algebraic approach in this section differs from that of Ruelle [Rue69], in that the functions are defined on subsets rather than subsequences. This is due to the fact that the model we are working on is defined readily on subsets of some finite set, rather than sequences of points in* $\mathbb{R}^3$*. It follows mainly the approach taken my Ueltschi and Poghosyan.*

We emphasise here that we are generalising to the case of complex valued polymer activities in order to use the context of Banach spaces in a later section. The physical model requires positive bounded polymer activities in order to have probabilities and to make sense physically.

**Definition** (Disjoint Union Notation)**.** *Throughout this section additive notation on sets is used to indicate a disjoint union.*

The mapping:

$$\Gamma : V_\Lambda^+ = \{\Phi | \Phi \in V_\Lambda \text{ and } \Phi(\emptyset) = 0\} \to \mathbb{1} + V_\Lambda^+ \tag{I.3.10}$$

called the $\star$-exponential, is defined by:

$$\Gamma\Phi = \mathbb{1} + \sum_{n=1}^{\infty} \frac{1}{n!}\Phi^{\star n} \tag{I.3.11}$$

$$(\Gamma\Phi)(X) = \sum_{X=\sum X_i} \prod_i \Phi(X_i) \tag{I.3.12}$$

Define the differentiation operation, for any subset $X \subset \Lambda$, by:

$$(D_X\Phi)(Y) := \Phi(X+Y)\delta_{\emptyset, X\cap Y} \tag{I.3.13}$$

The derivative has the following relations:

$$D_x(\Phi_1 \star \Phi_2)(Y) = ((D_x\Phi_1 \star \Phi_2)(Y) + (\Phi_1 \star D_x\Phi_2)(Y))\delta_{\emptyset,\{x\}\cap Y} \tag{I.3.14}$$

$$(D_x\Gamma\Phi)(Y) = (D_x\Phi \star \Gamma\Phi)(Y)\delta_{\emptyset,\{x\}\cap Y} \tag{I.3.15}$$

$$(D_X\Gamma\Phi)(Y) = \sum_{X=\sum_{i=1}^{q} X_1} (D_{X_1}\Phi \star \cdots \star D_{X_q}\Phi \star \Gamma\Phi)(Y)\delta_{\emptyset, X\cap Y} \tag{I.3.16}$$

Define:

$$(\lambda^N\Phi)(X) = \lambda^{N(X)}\Phi(X)\,\forall\lambda \in \mathbb{C} \tag{I.3.17}$$

Then, we have $\Gamma(\lambda^N\Phi) = \lambda^N\Gamma\Phi$.

For the generalisation of the subset gas to a complex polymer function $\Phi$, we may rewrite the following:

Taking $\Phi(\emptyset) = 0$, the partition function may be written as $\Xi_\Lambda = (\Gamma\Phi)(\Lambda)$.

In terms of this algebraic notation, the correlation functions (I.3.5) may be expressed as:

$$\rho_\Lambda(\gamma_1, \cdots \gamma_n) = \prod_{i=1}^{n} \Phi(\gamma_i) \frac{(\Gamma\Phi)(\Lambda \setminus \sum_{i=1}^{n} \gamma_i)}{(\Gamma\Phi)(\Lambda)} \text{ for } \gamma_i \cap \gamma_j = \emptyset \tag{I.3.18}$$

The reduced correlation functions, may be written as:

$$\bar{\rho}_\Lambda(X) = \frac{(\Gamma\Phi)(\Lambda \setminus X)}{(\Gamma\Phi)(\Lambda)} \tag{I.3.19}$$

where reduced correlation functions are defined by the following relation:

$$\rho_\Lambda(\gamma_1, \cdots \gamma_n) = \prod_{i=1}^{n} \Phi(\gamma_i)\bar{\rho}_\Lambda(\sum_i \gamma_i) \tag{I.3.20}$$

We have the sum rules:

$$\sum_{\gamma \ni x} \rho_\Lambda(\gamma) = 1 \tag{I.3.21}$$

$$\sum_{n=1}^{\infty} n\rho_\Lambda^{(n)} = \frac{1}{N(\Lambda)} \sum_{X \subset \Lambda} N(X)\rho_\Lambda(X) = 1 \tag{I.3.22}$$

These relations convey the fact that each site is occupied by one and only one particle. If we use equation (I.3.16), we obtain the relation:

$$\begin{aligned}(\Gamma\Phi)(Z+Y) &= (D_Y\Gamma\Phi)(Z) \\ &= (\Gamma\Phi)(Y)(\Gamma\Phi)(Z) + \sum_{Y=\sum Y_i} \sum_{\substack{S\subset Z \\ S\neq\emptyset}} (D_{Y_1}\Phi \star \cdots \star D_{Y_q}\Phi)(S)(\Gamma\Phi)(Z\setminus S)\end{aligned} \tag{I.3.23}$$

If we assume $(\Gamma\Phi)(\Lambda) \neq 0$ and rewrite $Z = \Lambda \setminus X$ with $Y \subset X \subset \Lambda$, we can then recast equation (I.3.23) (after dividing by $(\Gamma\Phi)(\Lambda)$) as:

$$\bar\rho_\Lambda(X\setminus Y) = (\Gamma\Phi)(Y)\bar\rho_\Lambda(X) + \sum_{Y=\sum Y_i} \sum_{\substack{S\subset\Lambda\setminus X \\ S\neq\emptyset}} (D_{Y_1}\Phi\star\cdots\star D_{Y_p}\Phi)(S)\bar\rho_\Lambda(X+S) \tag{I.3.24}$$

The correlation functions therefore satisfy $\forall X$ and $Y \subset X$ :

$$(\Gamma\Phi)(Y)\rho_\Lambda(X) = \chi_\Lambda(X)\bar\rho_\Lambda(X\setminus Y) - \sum_{Y=\sum_{i=1}^{q} Y_i} \sum_{S1\cdots S_q \atop \cup S_i = S\neq\emptyset \; S_i\cap S_j=\emptyset} \prod_{i=1}^{q} \Phi(Y_i+S_i)\bar\rho_\Lambda(X+S) \tag{I.3.25}$$

We have the following lemma, which reduces the number of equations we need to consider:

**Lemma I.3.1.** *If $\bar\rho_\Lambda$ satisfies equation (I.3.25) $\forall Y \subset X$ such that $N(Y) \leq n$, then it satisfies (I.3.25) $\forall Y$ such that $N(Y) = n+1$.*

**Corollary I.3.2.** *The following set of equations is equivalent to the full set (I.3.25):*

$$\Phi(x_1)\bar\rho_\Lambda(X) = \chi_\Lambda(X)\bar\rho_\Lambda(X\setminus\{x_1\}) - \sum_{\substack{S\neq\emptyset \\ S\subset\Lambda\setminus X}} \Phi(x_1+S)\bar\rho_\Lambda(X+S) \tag{I.3.26}$$

*for any $x_1 \in X$.*

This corollary gives precisely the *Kirkwood Salsburg equations*. Having an analytic solution to these equations provides us with an analytic equation for the

pressure introduced earlier.

In the homogeneous case, where the polymer function may be written as $\Phi(X) = z^{N(X)}$, we can see the connection to the classical gas cluster expansion. Finding the domain of convergence for the pressure is therefore connected to finding the domain of $z \in \mathbb{C}$ for which a solution to the (homogeneous) Kirkwood Salsburg equation exists.

### I.3.3 The Banach Space for the Kirkwood Salsburg Equations

We wish to apply this theory towards finding bounds on the cluster coefficients and understanding the analytic properties of the pressure function. This presentation of the Banach space setting for the Kirkwood Salsburg equation is influenced by the presentation in [Rue69]. In order to understand the set of $z \in \mathbb{C}$ for which the Kirkwood Salsburg equations have a unique solution, we must understand the setting of the equation in terms of an operator in a (complex) Banach Space. This version of the Banach Space is also the natural setting for Chapter IX.

We define $\mathcal{B}_\xi$ as the complex Banach space of complex functions $f$ on non-empty finite subsets of $\mathbb{Z}^\nu$, with the norm:

$$\|f\|_\xi = \sup_X \frac{|f(X)|}{\xi^{N(X)}} \tag{I.3.27}$$

where $\xi$ is positive.

For non-negative activity:

$$\bar{\rho}_\Lambda(X) \leq \frac{1}{\prod \Phi(x_i)} \tag{I.3.28}$$

For finite $\Lambda$ and nonnegative activity $\Phi$, $\bar{\rho}_\Lambda \in \mathcal{B}_\xi$ and satisfies $\|\bar{\rho}_\Lambda\|_\xi \leq 1$ if $\xi\Phi(X) \geq 1$ $\forall X$.

**Definition** (Kirkwood Salsburg Operator)**.** *We define the Kirkwood Salsburg operator $K_\xi$ for $f \in \mathcal{B}_\xi$:*

$$(K_\xi f)(x_1) = -\frac{1}{\Phi(x_1)} \sum_{\substack{S \neq \emptyset \\ x_1 \notin S}} \Phi(x_1 + S) f(x_1 + S)$$

$$(K_\xi f)(X) = \frac{1}{\Phi(x_1)} \left( f(X \setminus x_1) - \sum_{\substack{S \neq \emptyset \\ S \cap X = \emptyset}} \Phi(x_1 + S) f(X + S) \right) \; N(X) \geq 2 \tag{I.3.29}$$

**Definition** (The set $\Delta(\xi)$)**.** *We let $\Delta(\xi)$ be the set of complex activities $\Phi$ such that:*

$$\|\Phi\|_{\xi,\infty} = \sup_x \sum_{X \ni x} |\Phi(X)| \xi^{N(X)} < \infty \tag{I.3.30}$$

*and*

$$|\Phi(x)| > \sup_x \frac{1}{\xi} \left( 1 + \sum_{\substack{x \notin S \\ S \neq \emptyset}} |\Phi(x+S)| \xi^{N(x+S)} \right) = R(\xi) \tag{I.3.31}$$

Lemma I.3.3, below, indicates the characterisation of the domain of convergence for activity expansions.

**Lemma I.3.3.** *For any activity $\Phi$ such that $\sum_{X \ni x} \xi^{N(X)} |\Phi(X)| < \infty$ for some $\xi$ and such that $\min_X |\Phi(X)| \neq 0$, the operator $K_\Phi$ is a bounded operator on $\mathcal{B}_\xi$ and its norm satisfies:*

$$\|K_\Phi\|_\xi \leq \max \frac{1}{|\Phi(x)|} \frac{1}{\xi} \left( 1 + \sum_{\substack{x \notin S \\ S \neq \emptyset}} |\Phi(x+S)| \xi^{N(x+S)} \right) \tag{I.3.32}$$

*The operator $K_\Phi$ defined on the Banach space $\mathcal{B}_\xi$ is norm-analytic in $\Phi$ for $\Phi \in \Delta(\xi)$. In this domain $\|K_\Phi\|_\xi < 1$.*

**Theorem I.3.4** (Gruber and Kunz)**.** *For any $\Phi \in \Delta(\xi)$:*

*I) For any finite volume $\Lambda$, the Kirkwood Salsburg equation:*

$$\bar{\rho}_\Lambda = \chi_\Lambda \alpha + \chi_\Lambda K_\Phi \bar{\rho}_\Lambda \tag{I.3.33}$$

*has a unique solution in $\mathcal{B}_\xi$ obtained by iteration. The solution is norm analytic function of $\Phi$ in $\Delta(\xi)$, satisfying:*

$$\|\bar{\rho}_\Lambda\| \leq \left( 1 - \max_x \frac{1}{|\Phi(x)|\xi} \left( 1 + \sum_{S \neq \emptyset x \notin S} |\Phi(x+S)| \xi^{N(S)+1} \right) \right) \tag{I.3.34}$$

*II) For any finite volume $\Xi_\Lambda[\Phi] \neq 0$ and the solutions of the Kirkwood Salsburg equation coincide with the definitions of the correlation functions.*

*III) The infinite volume equation:*

$$\bar{\rho} = \alpha + K_\Phi \rho \tag{I.3.35}$$

*has a unique solution in $\mathcal{B}_\xi$, which is analytic in $\Delta(\xi)$. If the activity is invariant under a subgroup $T$ of translations then $\bar{\rho}$ is too.*

*IV) If the activity is invariant under a certain subgroup $T$ of translations such that the quotient group $\mathbb{Z}^\nu/T$ is finite, or if the activity has finite range, then there exists positive decreasing function $\varepsilon(\lambda)$ such that:*

$$\lim_{\lambda\to\infty} \varepsilon(\lambda) = 0 \tag{I.3.36}$$

*and*

$$|\bar{\rho}_\Lambda(X) - \bar{\rho}(X)| \leq \xi^{N(X)}\varepsilon(\lambda) \tag{I.3.37}$$

*where $\lambda$ is the minimum distance from $X$ to the boundary of $\Lambda$.*

*V) If the activity is invariant under a subgroup $T$ of translations, such that $\mathbb{Z}^\nu$ is finite, then:*

$$\bar{\rho}(X) = \lim_{\Lambda\to\infty} \frac{N(\mathbb{Z}^\nu/T)}{N(\Lambda)} \sum_{\tau in T} \bar{\rho}_\Lambda(\tau X) \tag{I.3.38}$$

*where $\Lambda \to \infty$ in the sense of van Hove. The convergence is uniform with respect to $\Phi(X)$ on every compact subset contained in $\Delta(\xi)$*

**Theorem I.3.5.** *The solutions $\bar{\rho}_\Lambda$ and $\bar{\rho}$ of the Kirkwood and Salsburg equation are analytic functions of the monomer activity $\Phi(x)$ in the domain:*

$$|\Phi(x)| > \max_x \frac{1}{\xi_0}\left(1 + \sum_{\substack{S\neq\emptyset \\ x\notin S}} |\Phi(x+S)|\xi_0^{N(S)+1}\right) = R(\xi_0) \tag{I.3.39}$$

*where $\xi_0$ is the value of $\xi$ at which $R(\xi)$ attains its minimum.*

*If the activity is translation invariant $\xi_0$ is the solution of:*

$$1 = \sum_{0\notin S} (N(S)-1)|\Phi(S)|\xi_0^{N(S)} \tag{I.3.40}$$

In the case where we let $\Phi(x) = z\,\forall x \in \mathbb{Z}^\nu$, then we are in the translation invariant case and have the following relation:

$$\lim_{\Lambda\to\infty} \frac{1}{N(\Lambda)} \sum_x \bar{\rho}_\Lambda(x) = \bar{\rho}^{(1)} = \frac{\rho^{(1)}}{z} \tag{I.3.41}$$

the convergence is uniform with respect to $z$ on every compact subset in the exterior of the circle of radius $R(\xi_0)$.

We obtain the following Mayer series for pressure and density:

$$\beta p = \ln z + \sum_{n=1}^{\infty} b_n z^{-n} \tag{I.3.42}$$

$$\rho^{(1)} = 1 + \sum_{n=1}^{\infty} n b_n z^{-n} \tag{I.3.43}$$

They have radius of convergence at least $R(\xi_0)^{-1}$ in $z^{-1}$.

## I.4 Polymer Models II - The Abstract Polymer Model

The paper of Kotecký and Preiss [KoPr86] indicated a movement away from understanding the cluster expansion convergence in terms of the coefficients. It gave abstract conditions to ensure the convergence of the series. This approach is informally called 'cluster expansions without expansions', which emphasises the technique of finding a condition for convergence, which makes no explicit reference to the coefficients or how the expansion appears. Further notable refinements of the Kotecký and Preiss bounds have been made by Dobrushin in [Dob96] and Fernández and Procacci in [FePr07]. Greater generalisation and interpretation of the Kotecký and Preiss conditions in various models of statistical mechanics alongside the use of further inequalities in the context of cluster expansion are used in the paper by Poghosyan and Ueltschi [PoUe09]. Furthermore, the conditions have been recast in terms of rooted tree fixed point equations and interpreted in this sense by Faris [Far10] and this is presented in Section VI.5.

In this section, the main idea behind the Kotecký and Preiss condition is presented as well as a remark upon the proof and directions for refinement. These have been achieved by the work of Dobrushin and Fernández and Procacci. This is given in Subsection I.4.2 and Section I.5.

### I.4.1 Notation

The abstract polymer model consists of:

i) A countable set $\mathfrak{P}$, whose elements are called polymers.

ii) A reflexive and symmetric relation $\iota \subset \mathfrak{P} \times \mathfrak{P}$, where we call $\gamma_1, \gamma_2 \in \mathfrak{P}$ incompatible if $(\gamma_1, \gamma_2) \in \iota$ and compatible if $(\gamma_1, \gamma_2) \notin \iota$.

We write $\gamma_1 \iota \gamma_2$ to denote an incompatible pair of polymers.

We denote for a subset $L \subset \mathfrak{P}$:

i) $\mathcal{B}(L)$ to be the family of finite subsets of $L$

ii) $\mathcal{D}(L)$ to be the subsets $\gamma \subset L$ consisting of mutually compatible polymers

iii) $\mathcal{D}_0(L)$ to be the finite subsets $\gamma \subset L$ consisting of mutually compatible polymers

We write $\mathcal{B} = \mathcal{B}(\mathfrak{P})$, $\mathcal{D} = \mathcal{D}(\mathfrak{P})$, and $\mathcal{D}_0 = \mathcal{D}_0(\mathfrak{P})$.

For $C \in \mathcal{B}$:

i) $|C|$ = number of polymers in $C$.

ii) $C\iota\gamma$ whenever $\exists\gamma' \in C$ such that $\gamma'\iota\gamma$

iii) $C$ is called a *cluster* if it is not decomposable into two (disjoint) non empty subsets $C = C_1 \cup C_2$, such that for all pairs $\gamma_1 \in C_1$ and $\gamma_2 \in C_2$, $(\gamma_1, \gamma_2) \in \iota$.

We define a decay rate $d : \mathfrak{P} \to [0, \infty)$. For every $C \in \mathcal{B}$, we define:

$$d(C) = \sum_{\gamma \in C} d(\gamma) \tag{I.4.1}$$

For our statistical weights we have a polymer functional: $\Phi : \mathfrak{P} \to \mathbb{C}$. For every $\partial \in \mathcal{D}_0$:

$$\Phi(\partial) = \prod_{\gamma \in \partial} \Phi(\gamma) \tag{I.4.2}$$

Under the proviso that $\Phi(\emptyset) = 1$
The partition function $\mathcal{Z}(L, \Phi)$ for all finite $L \subset \mathfrak{P}$ is:

$$\mathcal{Z}(L, \Phi) = \sum_{\gamma \in \mathcal{D}(L)} \Phi(\gamma) \tag{I.4.3}$$

To indicate the generality of the abstract polymer model, two main examples of an application to classical gas systems are given:

**Example** (Graph Weight Application)**.** *We observe the relationship with the classical gas here. If we define our countable set* $\mathfrak{P}$ *to be* $[n]^{(2)}$ *for the possible edges on our graph* $G$ *and take the diagonal subset of* $\mathfrak{P} \times \mathfrak{P}$ *as our* $\iota$*, we achieve:*

$$\mathcal{Z}([n]^{(2)}, \Phi) = \sum_{\gamma \in \mathcal{D}([n]^{(2)})} \Phi(\gamma) = \sum_{g \in \mathcal{G}[n]} \Phi(g) \tag{I.4.4}$$

*So we see that* $\Phi(g)$ *should correspond to our graph weights and hence the Kotecký and Preiss Criterion applies to this example.*

**Example** (Rewriting the Classical Gas Partition Function as a Polymer Partition Function [Pro07])**.** *In the classical gas, we have the potential function* $\Phi(x_i - x_j)$ *and the Boltzmann factor:*

$$e^{-\beta \sum\limits_{1\leq i<j\leq N} \Phi(x_i - x_j)} = \sum_{\{R_1,\cdots,R_s\}\in \mathsf{Par}[N]} \prod_{i=1}^{s} \xi(R_i) \tag{I.4.5}$$

*where* $\mathsf{Par}[N]$ *denotes the set of partitions on the set* $[N]$ *and we define* $\xi$ *by:*

$$\xi(R) := \begin{cases} 1 & \text{if } |R| = 1 \\ \sum\limits_{g\in\mathcal{C}[R]} \prod\limits_{\{i,j\}\in E(g)} (e^{-\beta\Phi(x_i-x_j)} - 1) & \text{if } |R| \geq 2 \end{cases} \tag{I.4.6}$$

*We then define:*

$$\zeta_{|R|} = \int_{\Lambda} \cdots \int_{\Lambda} \prod_{i\in R} \frac{\mathrm{d}x_i}{V} \xi(R) \tag{I.4.7}$$

*which is independent of the labels in the set. We thus can write the canonical partition function as a polymer partition function:*

$$Z_N(\Lambda,\beta) = \sum_{n\geq 0} \sum_{\substack{(R)_{[n]} \,|\, R_i\subset[N] \\ |R_i|\geq 2 \;\; R_i\cap R_j=\emptyset}} \zeta_{|R_1|}\cdots\zeta_{|R_n|} \tag{I.4.8}$$

*So the polymer set is the set of subsets of* $[N]$ *and two polymers are incompatible if they have non-empty intersection. The polymer weights are defined by the above.*

In order to remain on the same branch of the logarithm of $\mathcal{Z}$ for all small complex valued activities we need to make the following definition:

**Definition.** *We say that a topological space* $X$ *is* contractible *if the identity map on* $X$ *is null-homotopic, i.e. if it is homotopic to a constant map.*

Let $\mathfrak{M}$ be a contractible set of complex-valued polymer functionals. We assume that $\mathcal{Z}(L,\Phi) \neq 0 \;\forall \Phi \in \mathfrak{M}$. We may then uniquely define $\log \mathcal{Z}(L,\Phi)$ as the continuous branch of the logarithm for which:

$$\log \mathcal{Z}(L, \Phi = 0) = 0 \tag{I.4.9}$$

Note that $\mathcal{Z}(L,\Phi=0) = 1 \;\forall L \in \mathcal{B}$.

**Theorem I.4.1** (Kotecký-Preiss Criterion)**.** *If we have functions:* $a : \mathfrak{P} \to [0,\infty)$*;*

$d : \mathfrak{P} \to [0, \infty)$*; and* $\Phi : \mathfrak{P} \to \mathbb{C}$*, such that:*

$$\sum_{\gamma' | \gamma' \iota \gamma} e^{a(\gamma') + d(\gamma')} |\Phi(\gamma')| \leq a(\gamma) \tag{I.4.10}$$

*for each* $\gamma \in \mathfrak{P}$*. Then* $\mathcal{Z}(L, \Phi) \neq 0$ *for each finite* $L \subset \mathfrak{P}$ *and there exists a unique function* $\Phi^T : \mathcal{B} \to \mathbb{C}$ *such that:*

$$\log \mathcal{Z}(L, \Phi) = \sum_{C | C \subset L} \Phi^T(C) \tag{I.4.11}$$

*for every* $L \in \mathcal{B}$*. Moreover the function* $\Phi^T$ *is given by the formula:*

$$\Phi^T(C) = \sum_{B | B \subset C} (-1)^{|C| - |B|} \log \mathcal{Z}(B, \Phi) \tag{I.4.12}$$

*The estimate:*

$$\sum_{\substack{C \in \mathcal{B} \\ C \iota \gamma}} |\Phi^T(C)| e^{a(C)} \leq a(\gamma) \tag{I.4.13}$$

*holds true for every* $\gamma \in \mathfrak{P}$ *and* $\Phi^T(C) = 0$ *whenever* $C$ *is not a cluster.*

**Remark 4.** *The idea of finding the* $\Phi^T$ *is related to the notion of Möbius inversion on a partially ordered set. This is covered in greater detail in Chapter VII. Furthermore there are various interpretations of the form of this condition, as a fixed point equation, which are conveyed within Section VI.5*

The key place to look for improvements in such a bound is the overcounting that is used during the proof for the occurence of certain polymers incompatible with $\gamma$. This is addressed in the work of Fernández and Procacci [FePr07].

### I.4.2 Dobrushin's Criterion

A further refinement of the Kotecký Preiss criterion was made by Dobrushin in [Dob96]. He uses different notation to what has become standard in order to reduce confusion between the abstract polymer model and the physical idea of polymers as chains of particles, defined as subsets of the position space as done by Gruber and Kunz. In Dobrushin's notation the abstract polymers are called animals, clusters are herds and the (modified) $d$-function is interpreted as the 'might' of an animal. Furthermore, Dobrushin uses an underlying graph to determine the incompatibility relation, where two polymers are attached by an edge if they are incompatible. The

incompatibility relation, however, neglects self-edges. In this section, the incompatibility relation is understood to be the same as the previous section, except we are neglecting the diagonal. Dobrushin's paper also emphasises the use of a 'positive' polymer function/weight-function to find dominating values for the radii of the polydisc for convergence. This is instrumental in many of the proofs, including in the work of Fernández Procacci [FePr07]. The underlying polymer set in this case is finite and is represented by the notation $\mathfrak{F}$

**Theorem I.4.2** (Dobrushin's Criterion)**.** *Fix a positive weight function $\Phi_0 = \Phi_0(\gamma)$. For any polymer $\gamma \in \mathfrak{F}$:*

$$1 - \Phi_0(\gamma) \exp\left( \sum_{\substack{\gamma' \in \mathcal{F} \\ \gamma' \iota \gamma}} \Phi_0(\gamma') b(\gamma') \right) \geq \exp\left(-\Phi_0(\gamma) b(\gamma)\right) \tag{I.4.14}$$

*implies that*

$$\Phi_0(\gamma) \exp\left( \sum_{\substack{\gamma' \in \mathcal{F} \\ \gamma' \iota \gamma}} \Phi_0(\gamma') b(\gamma') \right) < 1 \tag{I.4.15}$$

*Let $W_0$ be the set of all weight functions $\Phi = \Phi(\gamma)$ $\gamma \in \mathfrak{F}$ such that $|\Phi(\gamma)| \leq \Phi_0(\gamma)$ $\forall \gamma \in \mathfrak{F}$.*

*For a finite set $L \subset \mathcal{F}$, the partition function $\mathcal{Z}(L; \Phi) \neq 0$ and for any finite $B \subset L$:*

$$\left| \ln \left| \frac{\mathcal{Z}(L; \Phi)}{\mathcal{Z}(B; \Phi)} \right| \right| \leq \sum_{\gamma \in L \setminus B} \Phi_0(\gamma) b(\gamma) \tag{I.4.16}$$

**Remark 5** (Connection to Kotecký Preiss Criterion)**.** *In the above notation, the Kotecký Preiss criterion may be written as:*

$$\exp\left( \sum_{\substack{\gamma' \in \mathfrak{F} \\ \gamma' \iota \gamma}} \Phi_0(\gamma') b(\gamma') + \Phi_0(\gamma) b(\gamma) \right) \leq b(\gamma) \tag{I.4.17}$$

*To put this precisely in the same form as in the previous section, we need to replace $b(\gamma)$ with $\exp(a(\gamma))$ and take the logarithm of both sides:*

$$\sum_{\substack{\gamma' \in \mathcal{F} \\ \gamma' \iota \gamma}} \Phi_0(\gamma') e^{a(\gamma')} + \Phi_0(\gamma) e^{a(\gamma)} \leq a(\gamma) \tag{I.4.18}$$

*In order to make a direct comparison between the two conditions, it is easiest to*

*consider the Kotecký Preiss criterion in the form of (I.4.17), which we rewrite as:*

$$\Phi_0(\gamma)\exp\left(\sum_{\substack{\gamma'\in\mathfrak{F}\\ \gamma'\iota\gamma}}\Phi_0(\gamma')b(\gamma')\right)\leq b(\gamma)\Phi_0(\gamma)\exp\left(-\Phi_0(\gamma)b(\gamma)\right) \tag{I.4.19}$$

*Dobrushin's criterion (I.4.14) may, in order to be more readily contrasted, be recast as:*

$$\Phi_0(\gamma)\exp\left(\sum_{\substack{\gamma'\in\mathfrak{F}\\ \gamma'\iota\gamma}}\Phi_0(\gamma')b(\gamma')\right)\leq 1-\exp\left(-\Phi_0(\gamma)b(\gamma)\right) \tag{I.4.20}$$

*If we let $x$ represent $\Phi_0(\gamma)b(\gamma)$, then comparing the strength of the inequalities amounts to comparing $xe^{-x}$ and $1-e^{-x}$ for positive $x$. The function $g(x) := 1-e^{-x}-xe^{-x}$ is positive for positive $x$, since $g(0)=1$ and $\frac{d}{dx}g(x)=xe^{-x}>0$ for positive $x$ and so it is increasing. The Dobrushin criterion therefore gives a larger radius of convergence than the Kotecký-Preiss condition.*

The proof of Dobrushin's criterion is inductive on the number of polymers. One can compute the ratio of the partition functions for a set polymers $\mathfrak{F}$ and a subset $\mathfrak{F}'\subset\mathfrak{F}$. One proceeds by removing a single polymer $\gamma\in\mathfrak{F}\setminus\mathfrak{F}'$. Letting $\tilde{\mathfrak{F}}=\mathfrak{F}\setminus\{\gamma\}$, then invoking the induction for the ratio of the partition functions for $\tilde{\mathfrak{F}}$ and $\mathfrak{F}'$ and using the Dobrushin condition to bound the final ratio of partition functions for $\mathcal{F}$ and $\tilde{\mathfrak{F}}$. This inductive approach is explicitly reorganised in the connections to the Kirkwood-Salsburg equations. The presentation of this proof has been further refined by the work of Miracle-Solé in [Mir00] and Bovier and Zahradník in [BoZa00].

## I.5 Fernández and Procacci Bounds

In the paper [FePr07], Fernández and Procacci present a new bound for cluster expansion on polymers. They give an in depth understanding of the nature of the Kotecký Preiss criterion and indicate where improvements can come in the case of hard-core interactions. Refinements are made by removing some of the excess terms used in the Kotecký Preiss bound. Further extensions and ideas are also presented in [BFP10]. The ideas in these papers can be used for both finite and countable sets, but we restrict our attention to finite sets of polymers $\mathfrak{F}$.

The model is a graph $g=(\mathfrak{F},E)$, which is called the interaction graph. Incompatibility is defined through the graph by $\gamma$ incompatible with $\gamma'$ if $\{\gamma,\gamma'\}\in E$.

This is denoted by writing $\gamma\iota\gamma'$. The edge set contains all edges from a vertex to itself, giving that $\iota$ is reflexive and symmetric. We have also a family of activities: $\boldsymbol{z} = \{z_\gamma\}_{\gamma\in\mathfrak{F}} \in \mathbb{C}^{\mathfrak{F}}$, replacing the polymer function $\Phi$, in order to ease notation. The probability distribution is:

$$W_{\mathfrak{F}}(\{\gamma_1, \cdots, \gamma_n\}) = \frac{1}{\Xi_{\mathfrak{F}}(z)} z_{\gamma_1} \cdots z_{\gamma_n} \prod_{j<k} \mathbb{1}_{\{\{\gamma_j,\gamma_k\}\notin E\}} \tag{I.5.1}$$

with:

$$\Xi_{\mathfrak{F}}(\boldsymbol{z}) = \sum_{n=0}^{\infty} \sum_{(\gamma_1,\cdots,\gamma_n)\in\mathfrak{F}^n} z_{\gamma_1} \cdots z_{\gamma_n} \prod_{i<j} \mathbb{1}_{\{\{\gamma_j,\gamma_i\}\notin E\}} \tag{I.5.2}$$

The logarithm and $\Phi^T$ are defined analogously to the previous section. We have the following formula for $\Phi^T$ ( coming from Cammarota [Cam82]):

$$\Phi^T(\gamma_1, \cdots, \gamma_n) = \begin{cases} 1 & \text{if } n = 1 \\ \sum_{G\in\mathcal{G}_{\{\gamma_1,\cdots,\gamma_n\}}} (-1)^{|E(G)|} & \text{if } n \geq 2 \text{ and } \mathcal{G}_{\{\gamma_1,\cdots,\gamma_n\}} \text{ connected} \\ 0 & \text{if } n \geq 2 \text{ and } \mathcal{G}_{\{\gamma_1,\cdots,\gamma_n\}} \text{ not connected} \end{cases} \tag{I.5.3}$$

The $\Phi^T$ are called Ursell functions and $\mathcal{G}_{\{\gamma_1,\cdots,\gamma_n\}}$ is the graph with vertex set $[n]$ and edge set $\{\{i,j\}|\gamma_i\iota\gamma_j\, 0 \leq i,j \leq n\}$.

The new criterion involves considering the modified partition function: $\Xi_{\Gamma(\gamma_0)}$ which is the partition function associated with restricting our attention to polymer functions on $\Gamma(\gamma_0) = \{\gamma \in \mathfrak{F}|\gamma\iota\gamma_0\}$. We make the more general definition that, for any finite family of polymers $X$:

$$\Gamma(X) := \{\gamma \in \mathfrak{F}|\exists\gamma' \in X \text{ such that } \gamma\iota\gamma'\} \tag{I.5.4}$$

This is called the neighbourhood of $X$.

We define the correlators in this theory as:

$$\phi_{\mathfrak{F}}(\mathbf{z};\gamma_1,\cdots,\gamma_p) = \left(z_{\gamma_1}\cdots z_{\gamma_p} \prod_{1\leq i<j\leq p} \mathbb{1}_{\{(\gamma_i,\gamma_j)\notin E\}}\right) \frac{\Xi_{\mathfrak{F}\setminus\Gamma(\cup_{i=1}^p\gamma_i)}(\mathbf{z})}{\Xi_{\mathfrak{F}}(\mathbf{z})} \tag{I.5.5}$$

The reduced correlators are:

$$\tilde{\phi}_{\mathfrak{F}}(\mathbf{z};X) := \frac{\Xi_{\mathfrak{F}\setminus X}(\mathbf{z})}{\Xi_{\mathfrak{F}}(\mathbf{z})} \tag{I.5.6}$$

The Mayer series is defined as:

$$\log \Xi_{\mathfrak{F}}(\mathbf{z}) = \sum_{i=1}^{\infty} \frac{1}{n!} \sum_{(\gamma_1,\cdots,\gamma_n)\in\mathfrak{F}^n} \phi^T(\gamma_1,\cdots\gamma_n) z_{\gamma_1}\cdots z_{\gamma_n} \tag{I.5.7}$$

We emphasise that we have the alternating sign property (proved in Section I.7):

$$|\phi^T(\gamma_1,\cdots,\gamma_n)| = (-1)^{n-1}\phi^T(\gamma_1,\cdots\gamma_n) \tag{I.5.8}$$

**Definition** (Auxiliary Series)**.** *Two related series for the logarithm of the partition function are:*

$$\begin{aligned}\Theta^{\mathfrak{F}}_{\gamma}(\mathbf{z}) :&= \log \Xi_{\mathfrak{F}}(\mathbf{z}) - \log \Xi_{\mathfrak{F}\setminus\{\gamma\}}(\mathbf{z}) \\ &= \sum_{n=1}^{\infty} \frac{1}{n!} \sum_{\substack{(\gamma_1,\cdots,\gamma_n)\in\mathfrak{F}^n \\ \exists i \text{ such that } \gamma_i=\gamma}} \phi^T(\gamma_1,\cdots\gamma_n) z_{\gamma_1}\cdots z_{\gamma_n} \end{aligned} \tag{I.5.9}$$

$$\Pi^{\mathfrak{F}}_{\gamma}(\mathbf{z}) := \frac{\partial}{\partial z_\gamma} \log \Xi_{\mathfrak{F}}(\mathbf{z}) \tag{I.5.10}$$

*The series (I.5.9) is related to the induction method used in Dobrushin's proof and the series (I.5.10) is called the pinned function.*

**Definition** (Positive Term Series)**.** *For a collection of positive numbers $\{\eta_\gamma\}_{\gamma\in\mathcal{F}}$, we define the positive term series:*

$$|\Theta|^{\mathfrak{F}}_{\gamma}(\boldsymbol{\eta}) := \sum_{n=1}^{\infty} \frac{1}{n!} \sum_{\substack{(\gamma_1,\cdots,\gamma_n)\in\mathfrak{F}^n \\ \exists i \text{ such that } \gamma_i=\gamma}} |\phi^T(\gamma_1,\cdots,\gamma_n)| \eta_{\gamma_1}\cdots\eta_{\gamma_n} \tag{I.5.11}$$

$$|\Pi|^{\mathfrak{F}}_{\gamma}(\boldsymbol{\eta}) := \sum_{n=0}^{\infty} \frac{1}{n!} \sum_{(\gamma_1,\cdots\gamma_n)\in\mathfrak{F}^n} |\phi^T(\gamma,\gamma_1,\cdots,\gamma_n)| \eta_{\gamma_1}\cdots\eta_{\gamma_n} \tag{I.5.12}$$

We dominate the auxiliary series term by term by the positive term series when $|z_\gamma| \le \eta_\gamma$. Convergence of the positive series thus implies absolute and uniform convergence for the auxiliary series in the polydisc:

$$\mathcal{D}_{\boldsymbol{\eta}} := \{\mathbf{z}| \, |z_\gamma| \le \eta_\gamma\} \tag{I.5.13}$$

**Theorem I.5.1** (Fernández and Procacci Criterion)**.** *If, for $\boldsymbol{\eta} \in [0,\infty)^{\mathfrak{F}}$, there exists $\boldsymbol{\nu} \in [0,\infty)^{\mathfrak{F}}$ such that*

$$\eta_{\gamma_0} \Xi_{\Gamma(\gamma_0)}(\boldsymbol{\nu}) \le \nu_{\gamma_0} \tag{I.5.14}$$

*for all $\gamma_0 \in \mathfrak{F}$, then the series $|\Theta|^{\mathfrak{F}}_{\gamma_0}(\boldsymbol{\eta})$ and $|\Pi|^{\mathfrak{F}}_{\gamma_0}(\boldsymbol{\eta})$ are convergent and satisfy:*

$$|\Pi|^{\mathfrak{F}}_{\gamma_0}(\boldsymbol{\eta}) \leq \Xi_{\Gamma(\gamma_0)}(\boldsymbol{\nu}) \tag{I.5.15}$$

$$|\Theta|^{\mathfrak{F}}_{\gamma_0}(\boldsymbol{\eta}) \leq -\ln(1-\eta_{\gamma_0})^{\Xi_{\Gamma(\gamma_0)}(\boldsymbol{\nu})-\nu_{\gamma_0}} \tag{I.5.16}$$

**Remark 6.** *We note that $\Xi_{\Gamma(\gamma_0)}(\boldsymbol{\eta}) \leq \prod_{\gamma \iota \gamma_0} (1+\eta_\gamma)$*

*The ways in which this is an improvement, comes from two alterations:*

*i) There are no monomials in $\Xi_{\Gamma(\gamma_0)}$ involving triangle diagrams in $\mathcal{G}$*

*ii) Only monomial containing $\eta_{\gamma_0}$ is $\eta_{\gamma_0}$.*

We can write our partition function as: $\Xi_{\Gamma(\gamma_0)}(\boldsymbol{\eta}) = \eta_{\gamma_0} + \Xi_{\Gamma^\star(\gamma_0)}(\boldsymbol{\eta})$, where $\Gamma^\star(\gamma_0) = \Gamma(\gamma_0) \setminus \{\gamma_0\}$.

**Corollary I.5.2** (Improved Dobrushin)**.** *If, for some $\boldsymbol{\eta} \in [0,\infty)^{\mathfrak{F}}$, there exists $\boldsymbol{\nu} \in [0,\infty)^{\mathfrak{F}}$ such that*

$$\eta_{\gamma_0}\left(\nu_{\gamma_0} + \prod_{\substack{\gamma \iota \gamma_0 \\ \gamma \neq \gamma_0}} (1+\nu_\gamma)\right) \leq \nu_{\gamma_0} \tag{I.5.17}$$

*for all $\gamma_0 \in \mathfrak{F}$, then $|\Theta|^{\mathfrak{F}}_{\gamma_0}(\boldsymbol{\eta})$ and $|\Pi|^{\mathfrak{F}}_{\gamma_0}(\boldsymbol{\eta})$ are convergent and satisfy:*

$$|\Pi|^{\mathfrak{F}}_{\gamma_0}(\boldsymbol{\eta}) \leq \prod_{\tilde{\gamma} \in \Gamma(\gamma_0)} (1+\nu_{\tilde{\gamma}}) \tag{I.5.18}$$

$$|\Theta|^{\mathfrak{F}}_{\gamma_0}(\boldsymbol{\eta}) \leq \ln(1+\nu_{\gamma_0}) \tag{I.5.19}$$

The three different criteria can be compared through the realisation that all of them are of the form:

$$\eta_{\gamma_0}\varphi_{\gamma_0}(\boldsymbol{\mu}) \leq \mu_{\gamma_0} \tag{I.5.20}$$

where the particular form of $\varphi$ is determined by the criterion:

$$\varphi_{\gamma_0}(\boldsymbol{\mu}) = \begin{cases} \exp\left(\sum_{\gamma \iota \gamma_0} \mu_\gamma\right) & \text{Kotecký-Preiss} \\ \prod_{\gamma \iota \gamma_0} (1+\mu_\gamma) & \text{Dobrushin} \\ \mu_{\gamma_0} + \prod_{\substack{\gamma \iota \gamma_0 \\ \gamma \neq \gamma_0}} (1+\mu_\gamma) & \text{Improved Dobrushin} \\ \Xi_{\Gamma(\gamma_0)}(\boldsymbol{\mu}) & \text{Fernández-Procacci} \end{cases} \tag{I.5.21}$$

**Remark 7** (Comments on the Comparison of these bounds)**.** *The Fernández-Procacci condition appears to be optimal from the point of view of including all of the information present in the incidence graph we are given. We use all of the information available to us from the particular vertex we are considering and its neighbours. Dobrushin's condition satisfies a simpler interpretation and is only dependent on what the neighbours of the chosen vertex are and it independent of how these vertices are connected to one another.*

In the example in Figure I.1, where $\gamma_0$ is the point about which we are understanding the function $\varphi$, in the improved Dobrushin case, we think of the partition function corresponding to the star graph in the figure, whereas the Fernández and Procacci bound uses the actual graph. For triangle free graphs this is the same.

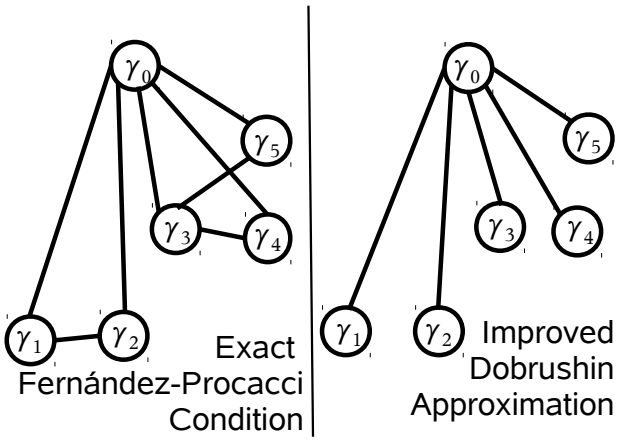



Figure I.1: The Incompatibility graph and Dobrushin's approximation

The proof of the Fernández-Procacci condition relies on the following:

**Theorem I.5.3** (Fernández-Procacci Condition as Fixed Point Equation)**.** *For fixed $\boldsymbol{\eta} \in [0,\infty)^{\mathfrak{F}}$, we define a map $T_{\boldsymbol{\eta}} : [0,\infty)^{\mathfrak{F}} \to [0,\infty]^{\mathfrak{F}}$, defined by:*

$$T_{\boldsymbol{\eta}}(\boldsymbol{\mu}) := \boldsymbol{\eta}\boldsymbol{\varphi}(\boldsymbol{\mu}) \tag{I.5.22}$$

*If we assume $\exists \boldsymbol{\mu} \in [0,\infty)^{\mathfrak{F}}$ satisfying:*

$$T_{\boldsymbol{\eta}}(\boldsymbol{\mu}) \leq \boldsymbol{\mu} \tag{I.5.23}$$

*Then:*

*i) $\exists \boldsymbol{\eta}^\star \in [0,\infty)^{\mathfrak{F}}$ such that $T^n_{\boldsymbol{\eta}}(\boldsymbol{\eta}) \uparrow \boldsymbol{\eta}^\star$ and $T_{\boldsymbol{\eta}}(\boldsymbol{\eta}^\star) = \boldsymbol{\eta}^\star$*

*ii) For each $n \in \mathbb{N}$:*

$$\boldsymbol{\eta\Pi} \leq \boldsymbol{\eta}^{\star} \leq T_{\boldsymbol{\eta}}^{n+1}(\boldsymbol{\mu}) \leq T_{\boldsymbol{\eta}}^{n}(\boldsymbol{\mu}) \leq \boldsymbol{\mu} \tag{I.5.24}$$

*for all of the above choices of $\boldsymbol{\varphi}$. The map $T_{\boldsymbol{\eta}}$ is monotonicity preserving and satisfies:*

$$\boldsymbol{\eta} \leq T_{\boldsymbol{\eta}}(\boldsymbol{\eta}) \leq T_{\boldsymbol{\eta}}(\boldsymbol{\mu}) \leq \boldsymbol{\mu} \tag{I.5.25}$$

The key point to this is the tree interpretation and the idea of a fixed point. This tree interpretation is used in Chapter VII.

### I.5.1 The Connection between Inductive Bounds and Kirkwood Salsburg Equations à la Gruber and Kunz

In the paper [BFP10], Bissacot, Fernández and Procacci present a useful interpretation of how the two different approaches of Dobrushin [Dob96] and Gruber and Kunz [GK71] are related. This section presents this connection. In this section, the collection of polymers $\mathfrak{K} := \mathcal{P}[\Lambda]$, where $\Lambda \subset \mathbb{Z}^n$ for some $n$. Subscripts of the thermodynamic functions are written in terms of the subset $\Lambda$ over which they are defined.

To make the connection to Section I.3, we have to understand how the abstract polymer conditions may be specialised to the subset gas. In this case the convergence criteria involve factorised weights of the form:

$$\nu_{\gamma} = \prod_{x \in \gamma} \xi_x =: \boldsymbol{\xi}^{\gamma} \tag{I.5.26}$$

for a family $\boldsymbol{\xi} = \{\xi_x\}_{x \in \Lambda}$, where we use the notation of Section I.3. Often it is convenient to write $e^{a_x} = \xi_x$ in this context. This is used in the two theorems below. Translating the two theorems into this context gives:

**Theorem I.5.4** (Dobrushin's Criterion for the Subset Gas)**.** *Let $a > 0$ and $\boldsymbol{\eta} = \{\eta_{\gamma}\}_{\gamma \in \mathfrak{K}}$ be collections of nonnegative numbers such that:*

$$\sup_{x \in \Lambda} \sum_{\substack{\gamma \in \mathfrak{K} \\ x \in \gamma}} \eta_{\gamma} e^{a|\gamma|} < e^{a} - 1 \tag{I.5.27}$$

*Then the function $|\Theta|_x^{\Lambda'}(\boldsymbol{\eta})$ is analytic on the polydisc $\mathcal{D}_{\boldsymbol{\eta}}$ and satisfies:*

$$|\Theta|_x^{\Lambda'}(\boldsymbol{\eta}) \leq \frac{1}{e^{a}} \left( 1 + \sup_{x \in \Lambda} \sum_{\substack{\gamma \in \mathfrak{K} \\ x \in \gamma}} \eta_{\gamma} e^{|\gamma|} \right) \tag{I.5.28}$$

*for all finite* $\Lambda' \subset \Lambda$ *and* $x \in \Lambda'$.

**Theorem I.5.5** (Fernández and Procacci Criterion for the Subset Gas)**.** *Let* $\mathbf{a} = \{a_x\}_{x\in\Lambda}$ *and* $\boldsymbol{\eta} = \{\eta_x\}_{x\in\Lambda}$ *be collections of nonegative numbers such that:*

$$\sum_{\substack{\gamma\in\Re\\ x\in\gamma}} \eta_\gamma e^{\sum\limits_{y\in\gamma} a_y} \leq e^{a_x} - 1 \tag{I.5.29}$$

*Then the function* $|\Theta|_x^{\Lambda'}(\boldsymbol{\eta})$ *is analytic on the polydisc* $\mathcal{D}_{\boldsymbol{\eta}}$ *and satisfies:*

$$|\Theta|_x^{\Lambda'}(\boldsymbol{\eta}) \leq a_x \tag{I.5.30}$$

*for all finite* $\Lambda' \subset \Lambda$ *and* $x \in \Lambda'$

*How Theorem I.5.1 implies Theorem I.5.5.* As described above, the connection needs to be made that:

$$\nu_\gamma = e^{\sum\limits_{x\in\gamma} a_x} \tag{I.5.31}$$

We write the Fernández and Procacci function as:

$$\varphi_\gamma^{\mathrm{FP}}(\boldsymbol{\nu}) = 1 + \sum_{n=1}^{|\gamma|} \frac{1}{n!} \sum_{(\gamma_1,\cdots\gamma_n)\in\Re^n} \prod_{i=1}^{n} \eta_{\gamma_i} e^{\sum\limits_{y\in\gamma_i} a_y} \prod_{1\leq i<j\leq n} \mathbb{1}_{\{\gamma_i\cap\gamma_j=\emptyset\}} \prod_{i=1}^{n} \mathbb{1}_{\{\gamma_i\cap\gamma\neq\emptyset\}} \tag{I.5.32}$$

A necessary condition for the indicator function to be satisfied is to have a particular $x_i \in \gamma$ to be associated with each $\gamma_i$. These $x_i$ should be distinct. This does not avoid the problems of whether they intersect outside of $\gamma$, however and so would give an upper bound for this positive function. We therefore have the inequality:

$$\varphi_\gamma^{\mathrm{FP}}(\boldsymbol{\nu}) \leq 1 + \sum_{n=1}^{|\gamma|} \frac{1}{n!} \sum_{\substack{(x_1,\cdots,x_n)\in\gamma^n\\ x_i\neq x_j \text{ for } i\neq j}} \prod_{i=1}^{n} \left( \sum_{\substack{\gamma\in\Re\\ x_i\in\gamma}} \eta_\gamma e^{\sum\limits_{y\in\gamma} a_y} \right) \tag{I.5.33}$$

If we apply the hypothesis (I.5.29), then we obtain:

$$\varphi_\gamma^{\mathrm{FP}}(\boldsymbol{\nu}) \leq 1 + \sum_{n=1}^{\|\gamma\|} \sum_{(x)_{[n]}\subset\gamma} \prod_{i=1}^{n} (e^{a_{x_i}} - 1) \tag{I.5.34}$$

$$= \prod_{x\in\gamma} ((e^{a_x} - 1) + 1) = e^{\sum\limits_{x\in\gamma} a_x} \tag{I.5.35}$$

This therefore gives us that $\eta_\gamma \varphi_\gamma^{\mathrm{FP}}(\boldsymbol{\nu}) \leq \eta_\gamma e^{\sum\limits_{x\in\gamma} a_x} = \nu_\gamma$ and so the Fernández and

Procacci condition (I.5.14) holds. We thus have (I.5.15): $|\Pi|_x^{\Lambda'}(\boldsymbol{\eta}) \leq \varphi_{\{x\}}^{\mathrm{FP}} \leq e^{a_x}$. Using the fact that in the subset gas model, $|\Pi|_{\{x\}}^{\Lambda}(\mathbf{z}) = \exp(|\Theta|_{\{x\}}^{\Lambda}(\mathbf{z}))$, we obtain the desired result. □

**Remark 8** (The connection between the method of Gruber and Kunz and the approach of Dobrushin)**.** *They key connection between the approach of Gruber and Kunz through the Kirkwood Salsburg equations and the inductive approach and Dobrushin, is that of the* site-addition identity*:*

$$\Xi_{Y\cup\{x\}}(\mathbf{z}) = \Xi_Y(\mathbf{z}) + \sum_{\substack{S\subset Y \\ |S|\geq 0}} z_{\{x\}\cup S}\Xi_{Y\setminus S}(\mathbf{z}) \tag{I.5.36}$$

*In the case of abstract polymers this is generalised to a* polymer-addition identity*:*

$$\Xi_{Z\cup\gamma_0}(\mathbf{z}) = \Xi_Z(\mathbf{z}) + z_{\gamma_0}\Xi_{Z\setminus\Gamma^\star(\gamma_0)}(\mathbf{z}) \tag{I.5.37}$$

*This is the essential ingredient of Dobrushin's proof.*

### I.5.2 Using Dobrushin's Induction Method for the Subset Gas

We indicate in this section how the inductive technique used by Dobrushin [Dob96], with refinements from Miracle-Solé [Mir00] may be applied to the subset gas to give the Fernández and Procacci condition (Theorem I.5.5).

**Proposition I.5.6** (Modified Fernández and Procacci)**.** *Let* $\mathbf{a} = \{a_x\}_{x\in\Lambda}$ *and* $\boldsymbol{\eta} = \{\eta_\gamma\}_{\gamma\in\mathcal{P}}$ *be collections of nonnegative numbers such that:*

$$\sum_{\substack{\gamma\in\Re \\ x\in\gamma}} \eta_\gamma e^{\sum\limits_{y\in\gamma} a_y} \leq e^{a_x} - 1 \tag{I.5.38}$$

*then:*

$$\frac{\Xi_{\Lambda'\setminus\{x\}}(-\boldsymbol{\eta})}{\Xi_{\Lambda'}(-\boldsymbol{\eta})} \leq e^{a_x} \tag{I.5.39}$$

*for any finite* $\Lambda' \subset \Lambda$ *and any* $x \in \Lambda'$*.*

This is proved using (I.5.36) and the reformulation of Dobrushin above.

In Section I.3.3, the Banach space is defined in terms of the parameter $\xi$. If all $\xi_x$ are the same in (I.5.26), then there is a straightforward correspondence. We could of course generalise to defining the Banach space as depending on $\boldsymbol{\xi} = \{\xi_x\}_{x\in\Lambda}$

rather than just a single parameter, giving the norm as:

$$\|f\|_{\boldsymbol{\xi}} := \sup_{X \in \mathfrak{K}} \frac{|f(X)|}{\boldsymbol{\xi}^X} \tag{I.5.40}$$

In this notation, we may rewrite (I.3.31) as:

$$\|K_{\mathbf{z}}\|_{\boldsymbol{\xi}} \leq \sup_{x \in \Lambda} \frac{1}{\xi_x} \left( 1 + \sup_{x \in \Lambda} \sum_{\substack{\gamma \in \mathfrak{K} \\ x \in \gamma}} |z_\gamma| \boldsymbol{\xi}^\gamma \right) \tag{I.5.41}$$

and use this to give a condition on $\mathbf{z}$ to have $\|K_{\mathbf{z}}\|_{\boldsymbol{\xi}} < 1$, to allow the inversion of the Kirkwood Salsburg equation.

If we change the variables so that $\xi_x = e^a$, making all variables the same, we obtain the Dobrushin condition (I.5.27):

$$\frac{1}{e^a} \left( 1 + \sup_{x \in \Lambda} \sum_{\substack{\gamma \in \mathfrak{K} \\ x \in \gamma}} \eta_\gamma e^{|\gamma| a} \right) < 1 \tag{I.5.42}$$

Hence in this case the two conditions are equivalent and so the inductive approaches can be seen in both examples.

## I.6 Interacting Polymer Model

In [Pro07], Procacci gives a polymer model with pairwise interactions. This is related to the model presented in the work of Suhov and Mazel [MaSu91]. The polymer set $\mathfrak{F}$ is finite. The potential energy function is:

$$\mathcal{U}(\gamma_1, \cdots, \gamma_n) = \sum_{1 \leq i < j \leq n} V(\gamma_i, \gamma_j) \tag{I.6.1}$$

where $V$ takes values in $\mathbb{R} \cup \{\infty\}$. In this case the measure is of the form:

$$\mathbb{P}(\gamma_1, \cdots, \gamma_n) = \frac{1}{\Xi_{\mathfrak{F}}} z_{\gamma_1} \cdots z_{\gamma_n} e^{-\sum_{1 \leq i < j \leq n} V(\gamma_i, \gamma_j)} \tag{I.6.2}$$

The interaction may encode an incompatibility condition: If $(\gamma, \gamma') \in \mathfrak{F} \times \mathfrak{F}$ is such that $V(\gamma, \gamma') = \infty$, then we write $\gamma \iota \gamma'$ and say the two polymers are incompatible. In the abstract polymer model setting, the stability condition is:

**Assumption 2** (Stability for Polymer Model)**.** *There exists a function* $B(\gamma) \geq 0$

*such that:*

$$\sum_{1\leq i<j\leq n} V(\gamma_i,\gamma_j) \geq -\sum_{i=1}^{n} B(\gamma_i) \tag{I.6.3}$$

*for all* $n$ *and all* $(\gamma_1,\cdots,\gamma_n)\in\mathfrak{F}^n$

If we have stability, then $\Xi_{\mathfrak{F}}$ converges and

$$\Xi_{\mathfrak{F}} \leq 1+\sum_{n\geq 1}\frac{1}{n!}\left(\sum_{\gamma\in\mathfrak{F}} z_\gamma e^{B(\gamma)}\right)^n \leq \exp\left(\sum_{\gamma\in\mathfrak{F}} z_\gamma e^{B(\gamma)}\right) \leq |\mathfrak{P}|\max_{\gamma\in\mathfrak{F}}\exp\left(z_\gamma e^{B(\gamma)}\right) \tag{I.6.4}$$

which means that $\Xi_{\mathfrak{F}}$ is analytic on $\mathbb{C}^{|\mathfrak{F}|}$.

In considering the Mayer expansion for the pressure, we can define a function $\Phi^T$, by Möbius inversion, so that:

$$\ln \Xi_{\mathcal{F}}(z) = \sum_{n=1}^{\infty}\frac{1}{n!}\sum_{(\gamma_1,\cdots,\gamma_n)\in\mathfrak{F}^n}\Phi^T(\gamma_1,\cdots,\gamma_n)z_{\gamma_1}\cdots z_{\gamma_n} \tag{I.6.5}$$

where we can define:

$$\Phi^T(\gamma_1,\cdots,\gamma_n) = \begin{cases} 1 & \text{if } n=1 \\ \sum\limits_{g\in\mathcal{C}[n]}\prod\limits_{\{i,j\}\in E(g)}\left(e^{-V(\gamma_i,\gamma_j)}-1\right) & \text{if } n\geq 2 \end{cases} \tag{I.6.6}$$

In order to consider the convergence of this expansion, we define a positive term sequence:

$$|\ln \Xi_{\mathfrak{F}}|(\boldsymbol{\eta}) = \sum_{n=1}^{\infty}\frac{1}{n!}\sum_{(\gamma_1,\cdots,\gamma_n)\in\mathfrak{F}^n}|\Phi^T(\gamma_1,\cdots,\gamma_n)|\eta_{\gamma_1}\cdots\eta_{\gamma_n} \tag{I.6.7}$$

We have that: $|\ln \Xi_{\mathfrak{F}}(z)| \leq |\ln \Xi_{\mathfrak{F}}|(\boldsymbol{\eta})$ and so if the latter series converges for a given $\boldsymbol{\eta}$, then the former series converges in the polydisc $\{|z_\gamma|\leq \eta_\gamma\}_{\gamma\in\mathfrak{F}}$.

We define the pinned sum:

$$\Pi_{\mathfrak{F}}^{\gamma_0}(\boldsymbol{\eta}) = \sum_{n=0}^{\infty}\frac{1}{n!}\sum_{(\gamma_1,\cdots,\gamma_n)\in\mathfrak{F}^n}|\Phi^T(\gamma_0,\gamma_1,\cdots,\gamma_n)|\eta_{\gamma_1}\cdots\eta_{\gamma_n} \tag{I.6.8}$$

and we have the inequality:

$$|\ln \Xi_{\mathfrak{F}}|(\boldsymbol{\eta}) \leq |\mathfrak{F}|\sup_{\gamma_0\in\mathfrak{F}}\eta_{\gamma_0}\Pi_{\mathcal{F}}^{\gamma_0}(\boldsymbol{\eta}) \tag{I.6.9}$$

**Theorem I.6.1** (Procacci - Kotecký Preiss condition for interacting polymers)**.** *Let* $\mu : \mathfrak{F} \to [0,\infty)\, \gamma \mapsto \mu_\gamma$ *be a non-negative valued function and assume that for each* $\gamma \in \mathfrak{F}$, $\exists \eta_\gamma \in [0,\infty)$ *such that:*

$$\eta_\gamma e^{B(\gamma)} \leq \mu_\gamma e^{-\sum\limits_{\tilde{\gamma}\in\mathfrak{F}} F(\gamma,\tilde{\gamma})\mu_{\tilde{\gamma}}} \tag{I.6.10}$$

*where*

$$F(\gamma_i,\gamma_j) := \begin{cases} |e^{-V(\gamma_i,\gamma_j)} - 1| = 1 & \text{if } (\gamma_i,\gamma_j) \in E \\ |V(\gamma_i,\gamma_j)| & \text{otherwise} \end{cases} \tag{I.6.11}$$

*then* $\Pi_{\gamma_0}(\eta)$ *converges and satisfies:*

$$\eta_{\gamma_0}\Pi_{\gamma_0}(\eta) \leq \mu_{\gamma_0} \tag{I.6.12}$$

**Remark 9.** *The key point of this theorem is that the hardcore abstract polymer model may have the Kotecký Preiss condition generalised to the interacting case. What is problematic here is that in the purely hardcore case this condition reduces to Kotecký Preiss rather than an improved Dobrushin or Fernández Procacci criterion.*

## I.7 Non-negative Potentials

### I.7.1 The Theorems

If we restrict our attention to non-negative potentials $U(q) \geq 0\, \forall q$, the cluster expansion becomes simplified somewhat. Groeneveld [Gro62, Gro67] gives tight bounds for the cluster coefficients and a proof that the signs of the cluster coefficients $b_l$ alternate.

**Theorem I.7.1** (Alternating sign of $b_l$)**.** *For non-negative potentials* $U(q) \geq 0$*, we have that the cluster coefficients* $b_l$ *alternate in sign, more precisely,* $(-1)^{l+1} b_l \geq 0$*.*

**Remark 10** (Non-Physicality of the Radius of Convergence)**.** *The fact that the cluster expansion has alternating signs for non-negative potentials determines that the point of non-analyticity closest to the origin for the pressure function in terms of fugacity must lie on the negative real axis. This makes it easier to apply particular techniques in searching for the radius of convergence of the series. It also implies that the point at which the cluster expansion is no longer valid need not be a phase transition, since the corresponding fugacity value will be 'unphysical'.*

**Theorem I.7.2** (Groeneveld's Bounds on Cluster Coefficients for Non-negative Potentials)**.** *For purely repulsive pair potentials in a classical gas, we have the following*

*bounds on the cluster coefficients.*

$$\frac{|2b_2|^{l-1}}{l} \leq |b_l| \leq \frac{|2b_2|^{l-1}l^{l-2}}{l!} \tag{I.7.1}$$

**Corollary I.7.3** (Groeneveld's Bounds on the Radius of Convergence for the Cluster Expansion for Non-negative potentials)**.** *The above upper and lower bounds for the coefficients in the cluster expansion, provide us with upper and lower bounds on the radius of convergence of the cluster expansion. We have, using the Cauchy Hadamard theorem:*

$$R^U = \lim_{l\to\infty}(|b_l^L|)^{-\frac{1}{l}} = \lim_{l\to\infty}\left(\frac{(2|b_2|)^{l-1}}{l}\right)^{-\frac{1}{l}} = \frac{1}{2|b_2|} \tag{I.7.2}$$

$$R^L = \lim_{l\to\infty}(|b_l^U|)^{-\frac{1}{l}} = \frac{1}{2|b_2|}\lim_{l\to\infty}\left(\frac{l!}{l^{l-2}}\right)^{\frac{1}{l}} = \frac{1}{2e|b_2|} \tag{I.7.3}$$

These bounds are Groeneveld's bounds in [Gro62]. These bounds are tight, in that there are two examples of positive potentials, where the coefficients are precisely the upper and lower bounds. These are found in Chapter III.

The proof of both theorems comes from a combinatorial identity expressed in terms of Mayer-weighted generating functions of particular types of graphs. What is presented below is an adaptation of what is found in the book of McCoy [McCoy10]

### I.7.2 The Combinatorial Identity

First some key graph definitions are given to understand the key structures.

**Definition** (Articulation Point)**.** *An articulation point in a connected graph $g$ is a vertex $v \in V(g)$ such that when we remove it and all the edges incident to it, our graph becomes disconnected.*

**Definition** (Ursell Functions)**.** *The Ursell function of $k$ positions $U_k(1,\cdots,k)$, maps $U : \mathbb{R}^{3k} \to \mathbb{R}$ and is defined as:*

$$U_k(1,\cdots.k) := \sum_{g\in\mathcal{C}[k]}\prod_{e\in E(g)} f_e \tag{I.7.4}$$

**Definition** ($D$-graphs)**.** *Define $\mathcal{D}[k]$ as the collection of connected graphs where 1 is not an articulation point. The related function $D_k(1,\cdots,k)$, $D : \mathbb{R}^{3k} \to \mathbb{R}$ is defined by:*

$$D_k(1,\cdots,k) := \sum_{g\in\mathcal{D}[k]}\prod_{e\in E(g)} f_e \tag{I.7.5}$$

We can write $D_{l-1}(1;2,\cdots,l)$ in terms of $U_{l-1}(2,\cdots l)$ with factors corresponding to connections from 1, since any graph in $\mathcal{D}[l]$ has all vertices in $[2,l]$ connected, when 1 is not present. We therefore have to understand adding in all possibilities of a non empty set of edges from 1. This amounts to for each edge deciding whether to add it or not, giving a $(1+f_{1,j})$-factor and then removing the possibility of adding none by taking away 1. This is independent of whether the potential is non-negative.

$$D_{l-1}(1;2,\cdots,l) = \left(\prod_{j=2}^{l}(1+f_{1,j})-1\right) U_{l-1}(2,\cdots,l) \tag{I.7.6}$$

$$= \left(\exp\left(-\beta\sum_{j=2}^{l}u(q_1-q_j)\right)-1\right) U_{l-1}(2,\cdots,l) \tag{I.7.7}$$

We also have the recursion relation:

$$\begin{aligned} U_l(1,\cdots,l) =& D_{l-1}(1;2,\cdots,l)U_1(1) + \sum_{3\le j\le l} D_{l-2}(1,2,\cdots,\hat{j},\cdots,l)U_2(1,j) \\ &+ \sum_{3\le j_1<j_2\le l} D_{l-3}(1;,2,\cdots,\hat{j}_1,\cdots,\hat{j}_2,\cdots l)U_3(1,j_1,j_2) \\ &+ \cdots + D_1(1;2)U_l(1,3,\cdots,l) \end{aligned} \tag{I.7.8}$$

This is explained in Section II.2.

**Remark 11** (The Combinatorial Structure)**.** *The combinatorial structure used here is that of the identity $\mathcal{B}'(\mathcal{C}^\bullet) = \mathcal{P}_+ \times \mathcal{C}$, which is explained in Chapter V. On the left hand side of the identity, we identify a particular point, in this case, that labelled* 1*. The right hand side considers the collection of connected graphs formed by removing the special point and then attaching them to this point, so that each connected component has at least one connection.*

### I.7.3 The Alternating Sign Property

For a nonnegative potential we have:

$$0 \le f_{i,j}+1 = \exp\left(-\beta u(q_i-q_j)\right) \le 1 \tag{I.7.9}$$

and therefore:

$$\prod_{j=2}^{l}(1+f_{1,j})-1 \le 0 \tag{I.7.10}$$

Hence $D_l$ and $U_l$ have opposite signs, from (I.7.7).

*Proof of Alternating sign Property.* We prove the alternating sign property by induction. For the base case, we note that:

$$U_1(1) = 1 > 0 \tag{I.7.11}$$

$$U_2(1,2) = f_{1,2} \leq 0 \tag{I.7.12}$$

$$D_1(1;2) = f_{1,2} \leq 0 \tag{I.7.13}$$

Take the induction hypothesis: for an integer $L$, we have for $l \leq L-1$:

$$(-1)^{l-1} U_l(1, \cdots, l) \geq 0 \tag{I.7.14}$$

$$(-1)^{l} D_l(1; 2, \cdots, l) \geq 0 \tag{I.7.15}$$

Therefore the sign of $D_{L-k}U_k$ is $(-1)^{L-1}$ and so from (I.7.8), the sign of $U_L$ is $(-1)^{L-1}$ and thus the sign of $D_L$ is $(-1)^L$ and hence the induction hypothesis holds for $L$. The sign of $b_k(V)$ is the same as the Ursell function $U_k(1, \cdots, k)$, hence we have:

$$0 \leq (-1)^{k-1} b_k(V) \leq (-1)^{k-1} b_k \tag{I.7.16}$$

□

**Remark 12.** *We note here that with graph-tree equations we can actually derive this result easily for nonnegative potentials. Since trees have a fixed number of edges and each edge has a negative weight we immediately get that the sign of $b_l$ is $(-1)^{l-1}$, but this involves understanding how weights have to be altered to create the graph-tree identity.*

### I.7.4 Upper and Lower Bounds

Using the notation:

$$f(1; 2, \cdots, l) = \prod_{j=2}^{l} (1 + f_{1,j}) - 1 \tag{I.7.17}$$

(I.7.7) can be written as:

$$D_l(1; 2, \cdots, l+1) = f(1; 2, \cdots, l+1) U_l(2, \cdots, l+1) \tag{I.7.18}$$

We note that:

$$\prod_{k=2}^{l}(1+f_{1,k}) = (1+f_{1,l})\prod_{k=2}^{l-1}(1+f_{1,k}) = f_{1,l}\prod_{k=2}^{l-1}(1+f_{1,k}) + \prod_{k=2}^{l-1}(1+f_{i,k}) \tag{I.7.19}$$

Therefore by induction:

$$\prod_{k=2}^{l}(1+f_{1,k}) = \sum_{k=3}^{l} f_{1,k}\prod_{j=2}^{k-1}(1+f_{1,j}) + (1+f_{1,2}) \tag{I.7.20}$$

And thus we have the equation:

$$f(1;2,\cdots,l) = f_{1,2} + \sum_{k=3}^{l} f_{1,k}\prod_{j=2}^{k-1}(1+f_{1,j}) \tag{I.7.21}$$

Take absolute value of this expression and use the triangle inequality to obtain:

$$|f(1;2,\cdots,l)| \leq |f_{1,2}| + \sum_{k=3}^{l} |f_{1,k}|\prod_{j=2}^{k-1}|1+f_{1,j}| \tag{I.7.22}$$

For nonnegative potentials $|1+f_{i,j}| \leq 1$, and so we obtain the bounds:

$$|f_{1,2}| \leq |f(1;2,\cdots,l)| \leq \sum_{k=2}^{l}|f_{1,k}| \tag{I.7.23}$$

If we integrate (I.7.18) over the coordinates $q_2,\cdots q_{l+1}$, we achieve:

$$|l!d_l| = \left|\int f(1;2,\cdots,l+1)U_l(2,\cdots,l+1)\,\mathrm{d}^D q_2\cdots\mathrm{d}^D q_{l+1}\right| \tag{I.7.24}$$

$$= \int |f(1;2,\cdots,l+1)||U_l(2,\cdots,l+1)|\,\mathrm{d}^D q_2\cdots\mathrm{d}^D q_{l+1} \tag{I.7.25}$$

The equality is from the signs we have for $U_l$ and $D_l$. Using the bounds on $|f(1;2,\cdots,l+1)|$, from (I.7.23), we obtain:

$$\int |f_{1,2}||U_l(2,\cdots,l+1)|\,\mathrm{d}^D q_2\cdots\mathrm{d}^D q_{l+1} \leq |l!d_l| \tag{I.7.26}$$

$$|l!d_l| \leq \int \sum_{k=2}^{l+1}|f_{1,k}||U_l(2,\cdots,l+1)|\,\mathrm{d}^D q_2\cdots\mathrm{d}^D q_{l+1} \tag{I.7.27}$$

Our functions depend only on the differences between the $q$ variables they contain and so the integrals are independent of one of the integration variables. [See Remark

1 in Section I.2] The functions $f_{1,k}$ and $U_l(2,\cdots,l+1)$ have only the point $k$ in common and so we can separate the two integrals, by neglecting the $k$ integral for $U$. We also note that in the upper bound the integrals do not depend on the label and so we can replace the sum by $l|f_{1,2}|$. We then get the expression:

$$\int |f_{1,2}|\,\mathrm{d}^D q_2 \int |u_l(2,\cdots,l+1)|\,\mathrm{d}^D q_3 \cdots \mathrm{d}^D q_{l+1} \leq |l! d_l| \tag{I.7.28}$$

$$|l! d_l| \leq \int l|f_{1,2}|\,\mathrm{d}^D q_2 \int |U_l(2,\cdots,l+1)|\,\mathrm{d}^D q_3 \cdots \mathrm{d}^N q_{l+1} \tag{I.7.29}$$

We can use definitions for $b_l$ to obtain bounds for $|d_l|$:

$$2|b_2||b_l| \leq |d_l| \leq 2|b_2|l|b_l| \tag{I.7.30}$$

We can rewrite (I.7.8) in terms of modulus signs:

$$l(l-1)|b_l| = \sum_{k=1}^{l-1} k|b_k|(l-k)|d_{l-k}| \tag{I.7.31}$$

We relabel for $l \geq 1$:

$$l|b_l| = a_{l-1} \tag{I.7.32}$$

to get:

$$l a_l = \sum_{m=0}^{l-1} a_m (l-m)|d_{l-m}| \tag{I.7.33}$$

Since everything is positive upper bounds and lower bounds for $a_m$ can be achieved by substituting the upper and lower bounds for $t_l$ respectively. We define the quantities $a_l^U$ and $a_l^L$ as the solutions of the recursion relations:

$$l a_l^{U,L} = \sum_{m=0}^{l-1} a_m^{U,L} (l-m) d_{l-m}^{U,L} \tag{I.7.34}$$

with initial condition

$$a_1^L = a_1^U = a_1 = 2|b_2| = B \tag{I.7.35}$$

We have the upper and lower $d$-values, defined as:

$$d_l^L = B\frac{a_{l-1}^L}{l} \tag{I.7.36}$$

$$d_l^U = B a_{l-1}^U \tag{I.7.37}$$

We need to prove that these really do give the upper and lower bounds we claim they do. We prove this by induction. The base case for $l = 1$ is certainly true since they are all equal. So assume:

$$a_l^L \leq a_l \leq a_l^U \tag{I.7.38}$$

for $k \leq l - 1$ and thus use the recurrence relations so that:

$$la_l = \sum_{m=0}^{l-1} a_m(l-m)|t_{l-m}| \leq \sum_{m=0}^{l-1} a_m^U(l-m)d_{l-m}^U = la_l^U \tag{I.7.39}$$

and

$$la_l = \sum_{m=0}^{l-1} a_m(l-m)|t_{l-m}| \geq \sum_{m=o}^{l-1} a_m^L(l-m)d_{l-m}^L = la_l^L \tag{I.7.40}$$

and hence it is true for all $l$ by induction.

### I.7.5 Functional and Differential Equations for Upper and Lower bounds

We define the generating functions:

$$A^{U,L}(z) = \sum_{l=0}^{\infty} a_l^{U,L} z^l \tag{I.7.41}$$

$$D^{U,L}(z) = \sum_{l=0}^{\infty} d_l^{U,L} z^l \tag{I.7.42}$$

We have from the recurrence relations (I.7.34) that:

$$A^{U,L}(z) = \exp(D^{U,L}(z)) \tag{I.7.43}$$

We want to use the definitions for $d_l^{U,L}$ from the equations (I.7.36) and (I.7.37) in equation (I.7.42) to find a second relation to give a functional/differential equation for the $A^{U,L}(z)$.

For the upper bound we substitute $d^U$ for $a^U$, using (I.7.37), in (I.7.42) to obtain:

$$D^U(z) = BzA^U(z) \tag{I.7.44}$$

For the lower bound we substitute $d^L$ for $a^L$, using (I.7.36), into (I.7.42) to first obtain:

$$D^L(z) = \sum_{l=1}^{\infty} z^l \frac{B}{l} a_{l-1}^L \tag{I.7.45}$$

which gives us the differential equation:

$$\frac{\partial D^L(z)}{\partial z} = BA^L(z) \tag{I.7.46}$$

### I.7.6 The Lower Bound

From (I.7.43) we can calculate:

$$\frac{\partial A^L(z)}{\partial z} = A^L(z)\frac{\partial D^L(z)}{\partial z} \tag{I.7.47}$$

If we use (I.7.46) to substitute for the derivative of $D^L$, we get the differential equation:

$$\frac{\partial A^L(z)}{\partial z} = B(A^L(z))^2 \tag{I.7.48}$$

If we use the initial condition $A^L(0) = 1$, we obtain the equation:

$$A^L(z) = \frac{1}{1 - Bz} \tag{I.7.49}$$

**Remark 13.** *$A^L(z)$ is the generating series of permutations with weight $B^n$.*

We find that $a_l^L = B^l$ (since it is a geometric series) and so by definition:

$$|b_l^L| = \frac{|B^{l-1}|}{l} \tag{I.7.50}$$

Using the alternating signs property, we have the lower bound:

$$\frac{1}{l} \leq \frac{b_l}{|2b_2|^{l-1}} \tag{I.7.51}$$

### I.7.7 The Upper Bound

If we substitute (I.7.44) into (I.7.43), we obtain the functional equation:

$$A^U(z) = \exp(BzA^U(z)) \tag{I.7.52}$$

We can take the logarithm to obtain:

$$z = \frac{\ln A^U(z)}{BA^U(z)} \tag{I.7.53}$$

We can obtain the desired coeffcients $a_l^U$ by taking the contour integral:

$$a_l^U = \frac{1}{2\pi i}\int_C \mathrm{d}z\, \frac{A^U(z)}{z^{l+1}} \tag{I.7.54}$$

From our functional relation, we find that:

$$\mathrm{d}z = (1 - \ln A^U)\frac{\mathrm{d}A^U}{B(A^U)^2} \tag{I.7.55}$$

If we set $A^U = e^\xi$, we have:

$$z = \frac{\xi}{Be^\xi} \tag{I.7.56}$$

$$\mathrm{d}z = \frac{1-\xi}{Be^\xi}\mathrm{d}\xi \tag{I.7.57}$$

The closed contour $C$ enclosing 0 in the $z$-plane will map to a closed contour $C'$ enclosing 0 in the $\xi$-plane. Thus we find:

$$a_l^U = \frac{1}{2\pi i}B^l \int_{C'} \mathrm{d}\xi\, \frac{1-\xi}{\xi^{l+1}} e^{\xi(l+1)} \tag{I.7.58}$$

This can be computed to give:

$$a_l^U = \frac{B^l(l+1)^{l-1}}{l!} \tag{I.7.59}$$

Therefore with alternating sign property, we obtain the upper bound:

$$\frac{b_l}{|2b_2|^{l-1}} \leq \frac{l^{l-2}}{l!} \tag{I.7.60}$$

The final bounds for the coefficients are:

$$\frac{|2b_2|^{l-1}}{l} \leq |b_l| \leq \frac{|2b_2|^{l-1}l^{l-2}}{l!} \tag{I.7.61}$$

## I.8 Models and Context of Cluster Expansion

In [PoUe09], Poghosyan and Ueltschi give a general background to cluster expansions and the models in Statistical Mechanics they apply to. The general abstract setting is that of a set $\mathbb{X}$ with the structure of an abstract measure space: $(\mathbb{X}, \mathcal{X}, \mu)$, with $\mu$ a complex measure. We have two complex measurable symmetric functions on

$\mathbb{X} \times \mathbb{X}$: $\zeta$ and $u$, which are related by:

$$\zeta(x, y) = e^{-u(x,y)} - 1 \tag{I.8.1}$$

We allow the real part of $u$ to take the value $\infty$, in which case $\zeta(x, y) = -1$. $u(x, y)$ represents the interaction between a particle at $x$ and a particle at $y$ and $u(x, y) = \infty$ indicates a hard core repulsion.

This is a further generalisation of the abstract polymer model, where the complex measurable symmetric function $u$ replaces the relation. $\mathbb{X}$ replaces the set of polymers $\mathfrak{P}$ and $\mu$ is the analogue of the weights for the polymer function. This set up and the bounds attained from this are useful in Chapter IV in obtaining alternative bounds on the virial expansion coefficients and in Chapter IX, when considering the context in which the analysis of the Lagrange inversion may be applied.

For these abstract models we have the partition function defined by:

$$Z = \sum_{n \geq 0} \frac{1}{n!} \int \mathrm{d}\mu(x_1) \cdots \int \mathrm{d}\mu(x_n) \exp\left( - \sum_{1 \leq i < j \leq n} u(x_i, x_j) \right) \tag{I.8.2}$$

which we can write in terms of the $\zeta$'s:

$$Z = \sum_{n \geq 0} \frac{1}{n!} \int \mathrm{d}\mu(x_1) \cdots \int \mathrm{d}\mu(x_n) \prod_{1 \leq i < j \leq n} (1 + \zeta(x_i, x_j)) \tag{I.8.3}$$

**Assumption 3** (Stability)**.** *In this general abstract setting the relevance of the d-function from the Kotecký-Preiss condition is related to the notion of stability for our potentials. We assume the existence of a non-negative function b on $\mathbb{X}$ such that for all $n$ and almost all $x_1, \cdots, x_n \in \mathbb{X}$:*

$$\prod_{1 \leq i < j \leq n} |1 + \zeta(x_i, x_j)| \leq \prod_{i=1}^{n} e^{b(x_i)} \tag{I.8.4}$$

*Which means a lower bound for $\Re u$:*

$$\sum_{1 \leq i < j \leq n} \Re(u(x_i, x_j)) \geq - \sum_{i=1}^{n} b(x_i) \tag{I.8.5}$$

**Assumption 4** (Kotecký and Preiss Criterion)**.** *The generalisation of the Kotecký-*

*Preiss condition (I.4.10) to the abstract setting is thus:*

$$\int \mathrm{d}|\mu|(y)\,|\zeta(x,y)|e^{a(y)+2b(y)} \leq a(x) \tag{I.8.6}$$

*where $|\mu|$ denotes the total variation of the measure $\mu$.*

If we define the modified potential:

$$\bar{u}(x,y) = \begin{cases} u(x,y) & \text{if } \Re(u(x,y)) \neq \infty \\ 1 & \text{if } \Re(u(x,y)) = \infty \end{cases} \tag{I.8.7}$$

**Assumption 5** (Procacci's Variation of Kotecký and Preiss Criterion)**.** *In [Pro07] a slightly different version of the Kotecký and Preiss criterion is given as: There exists a non-negative function $a$ on $\mathbb{X}$ auch that for almost all $x \in \mathbb{X}$:*

$$\int \mathrm{d}|\mu|(y)\,|\bar{u}(x,y)|e^{a(y)+b(y)} \leq a(x) \tag{I.8.8}$$

The corresponding connected function or Ursell Function for this model is:

$$\phi(x_1,\cdots,x_n) = \begin{cases} 1 & \text{if } n=1 \\ \frac{1}{n!}\sum\limits_{g\in\mathcal{C}[n]}\prod\limits_{\{i,j\}\in E(g)}\zeta(x_i,x_j) & \text{if } n\geq 2 \end{cases} \tag{I.8.9}$$

The cluster expansion Theorem for a general abstract model is:

**Theorem I.8.1** (Cluster Expansions)**.** *Assuming I.8.4 and I.8.6 or I.8.4 and I.8.8, and that $\int \mathrm{d}|\mu|(y)\,e^{a(y)+2b(y)} < \infty$, then we have:*

$$Z = \exp\left(\sum_{n\geq 1}\int \mathrm{d}\mu(x_1)\cdots\mathrm{d}\mu(x_n)\,\phi(x_1,\cdots,x_n)\right) \tag{I.8.10}$$

*The term in the exponential converges absolutely. Furthermore, for almost all $x_1 \in \mathbb{X}$, we have the following estimate:*

$$\sum_{n\geq 2} n\int \mathrm{d}|\mu|(x_2)\cdots\int \mathrm{d}|\mu|(x_n)\,|\phi(x_1,\cdots,x_n)| \leq (e^{a(x_1)}-1)e^{2b(x_1)} \tag{I.8.11}$$

The modified bounds from the Procacci condition [Pro07] are used in Section IV.6 to achieve improvements on virial expansion bounds.

## Conclusions & Open Questions

Groeneveld has provided optimal bounds for the cluster expansion for positive potentials as:

$$\frac{|2b_2|^{l-1}}{l} \leq b_l \leq \frac{|2b_2|^{l-1} l^{l-2}}{l!} \tag{I.8.12}$$

Extensions have been given towards potentials assuming only stability. The key approach in this case is to take Poghosyan and Ueltschi's [PoUe09] formulation of the Kotecký Preiss condition, combining the Tree-Graph identity of Brydges Battle Federbush and Kennedy [BaFe84, Bry84, BrFe78, BrKe87] (covered in Chapter VII) into an alternative bound, which is comparable to Procacci [Pro07].

The results of this chapter offer a useful starting point to studying bounds for the virial expansion covered in the following chapters. Furthermore, other techniques of cluster expansions are explored and adapted towards the virial expansions. The tree interpretations of the methods of bounding cluster expansions in this chapter provide suitable bounds, but it would be helpful to see the coefficients resolved for further particular models and a clear connection between all of the different approaches which is instigated by this chapter. Furthermore, recent work by Leroux and collaborators [BLL98, DLL07, Ler04] and Faris [Far08, Far10] has indicated the connection of Joyal's [Joy81] Combinatorial Species of Structure to statistical Mechanics. These connections are introduced in Chapters V and VI and in particular extended throughout this thesis.

# Chapter II

# The Virial Expansion

This chapter is a background review of various approaches taken to understand the convergence and coefficients of the virial expansion. The van der Waals Equation of State was the first approximation for a virial expansion. The chapter covers Mayer's proof of the connection between virial coefficients and weighted two-connected graphs. This is called Mayer's second theorem. The method of Lagrange inversion on the cluster expansion to achieve virial expansion bounds through the Kirkwood Salsburg equations is presented to motivate Chapter IV. This chapter presents also the approach of Pulvirenti and Tsagkarogiannis [PT12] through the Canonical Ensemble, which avoids the unphysical singularity present in the cluster expansion.

The word 'virial' comes from the Latin for vital force or energy. It is a term coined by the German Mathematical Physicist, Rudolf Julius Emmanuel Clausius (1822-1888). It was used to describe how the average kinetic and potential energies are related to the total system energy and internuclear forces. The idea of a virial expansion is that external properties of a system - the pressure and density - are related to each other through the internal forces - the pair potential. The virial expansion was introduced in 1885 by Thiessen [Thi85] and developed by Heike Kamerlingh-Onnes in his papers [Kam01] and [Kam02]. The virial expansion is for low densities and generally high temperatures.

The cluster expansion is an expansion in activity $z = e^{\beta\mu}$, where $\mu$ is the chemical potential. If we wish to use density $\rho$ as our order parameter instead, then we must use the Legendre transform. Density, $\rho$, is the conjugate variable to chemical potential, $\mu$, in this sense. The Legendre transform from chemical potential to density, transforms the pressure function into a free energy function.

The van der Waals Equation of State was proposed by Johannes Diderik

van der Waals in 1873, but suffers from certain problems at the liquid-gas phase transition. James Clerk-Maxwell in 1875 proposed an equal areas rule to the curve we obtain from the van der Waals Equation of State. This modification to the van der Waals virial expansion corrects the issue of pressure increasing as volume increases. The equal areas rule adds a straight line to this graph with equal area above it and below it in the original graph. This is equivalent to taking an appropriate convex envelope of the pressure function. The ideas are present in the short set of notes [Des13].

The first main systematic study of the virial expansion comes from Mayer [MMay40], in which an expansion is derived for the coefficients in terms of weighted two-connected graphs. The main assumption is that the potential can be written as a sum of two-body potentials. These ideas were inspired and are related to the cluster integrals of the previous chapter. One finds that the cluster integrals can be expressed as products and sums of what are called irreducible integrals, which are precisely the integrals corresponding to two-connected graphs in the Mayer expansion of the previous section. The method of obtaining such expressions is compared with an approach involving the use of Joyal's combinatorial species of structure [Joy81]. The techniques of combinatorial species of structure, which are discussed in the Chapter V, make the somewhat cumbersome calculations made by Mayer a lot neater and more powerful, in the sense that they have a greater generality under which they can be interpreted.

Often in the literature, to understand the convergence of cluster and virial expansions, 'pinned sums' or (reduced) correlation functions are introduced and we have fugacity series for these. The use of introducing the correlation functions, which are of greater generality than the density, is that we can get integral equation relations between the correlation functions. These may be viewed as fixed point equations in a suitable Banach Space. This is the approach of Kirkwood and Salsburg equations [KiSa53] and the further generalisations [LePer63]. The application of these to understanding the virial expansion is twofold: firstly, we directly get suitable bounds on density and fugacity, which can be used in an inversion relationship for pressure in terms of density; secondly, they also provide access to important inequalities explained in [LePer63]. Similar inequalities are found in the papers by Penrose [Pen63b] and Lieb [Lieb63]. They offer bounds on the radius of convergence of both the cluster and virial expansion and are particularly useful for positive potentials, although there is a simple generalisation to stable potentials.

Using the Canonical Ensemble, rather than the Grand Canonical Ensemble, avoids the cluster expansion as an intermediate stage. The cluster expansion inver-

sion often results in having to consider unphysical restrictions. For inversion one is required to remain in a domain of analyticity for $\beta P(z)$ for $z \in \mathbb{C}$. We note that for positive potentials the point of non analyticity occurs at negative real $z$, which is unphysical as this is the exponential of the real parameters $\beta$ and $\mu$. Canonical Ensemble calculations are made for the free energy, in which density is the order parameter of the model. This was undertaken by Pulvirenti and Tsagkarogiannis [PT12] and then further developed by Morais and Procacci [MoPr13]. These also involve understanding the relationship of Canonical Ensemble representations with Polymer Models, for which there have been many good results as detailed in Chapter I.

Most of the ideas in this chapter can also be understood from the point of view of combinatorial species of structure. Many of the relationships can be succinctly described by such tools and some derivations are more illuminating and straightforward in that context. The application of these combinatorial species of structure is remarked upon throughout this chapter, but Chapter V contains all of the details of these relationships and an understanding of this theory.

## II.1 The van der Waals Equation of State

In this section the van der Waals equation of state is presented and interpreted with reference to [MMay40, Hill56].

The first equation of state discovered experimentally is the ideal gas law:

$$VP = NkT \tag{II.1.1}$$

This is derived in Section I.1.1. The early history of virial expansions focused on making reasonable approximations, which can be tuned to experiments.

The van der Waals equation of state is a semi-empirical equation of state. It is derived rigorously in the paper [LePen66]. The ideal gas law is a good approximation for sparse gases, or gases at high temperatures, so that intermolecular forces are unimportant. However, in the situation where a real gas is compressed so that intermolecular forces become important, we need to find a reasonable approximation. The van der Waals equation of state provides this approximation for the pressure-volume relationship, even close to the condensed phase. The success of the van der Waals equation of state comes from its ability to be accurate up to second order in the density and that for relatively small densities the higher order terms are irrelevant. The van der Waals equation is:

$$\left(P + \frac{N^2 a}{V^2}\right)(V - Nb) = NkT \tag{II.1.2}$$

The $b$ corresponds to the volume exclusion effect of the particles. This is the simple approximation that particles can't occupy the same space. It corresponds to a repulsive close range interaction in the potential. If we consider the dimensions of $b$, we realise, since $V$ is in the same bracket it must somehow be proportional to a volume. There are $N$ particles so it corresponds to the excluded volume per particle. This gives the perturbation in $\rho$ on the RHS of the ideal gas law (II.1.1).

In considering the perturbation in pressure, caused by the interactions, we realise that there are approximately $N^2$ pairwise interactions and we write $a$ to be the integral over the potential:

$$\int U(x)\mathrm{d}^3x \tag{II.1.3}$$

In order to have an expression with the correct dimensions for pressure here, we are required to divide by $V^2$. The dimensions of $a$, interpreted as the integral of the potential are $ML^5T^{-2}$ and the dimensions of pressure are $ML^{-1}T^{-2}$ so we have a discrepancy of $L^6$ which is accounted for by the division by the squared volume term.

We may expand equation (II.1.2) to put it in the form:

$$PV + \frac{N^2 a}{V} - PNb - \frac{N^3 ab}{V^2} = NkT \tag{II.1.4}$$

and rearrange to get:

$$\beta P(1 - b\rho) = \rho - a\beta\rho^2 + ab\beta\rho^3 \tag{II.1.5}$$

where $\rho = \frac{N}{V}$.

If we divide by $(1 - b\rho)$ on both sides we achieve something resembling a virial expansion:

$$\beta P = \rho(1 - a\beta\rho(1 - b\rho)) \sum_{k=0}^{\infty} b^k \rho^k \tag{II.1.6}$$

This has first term $\rho$, as the ideal gas law, as we would expect. The second term is $(b - a\beta)\rho^2$, which we will proceed to illuminate. This van der Waals Equation can therefore give the exact behaviour of systems up to the second order and this is to what it owes its success for low density systems. The other terms are given as $b^{k-1}\rho^k$. This geometric expression for the virial coefficients reflects the geometric bounds that are often found for virial coefficients. Many bounds achieved for the

virial expansion, involve the coefficients being bounded in the form:

$$|c_k| \leq f(\beta)G(\beta)^{k-1} \tag{II.1.7}$$

In this sense the power of the van der Waals equations relies on the fact that we can achieve (relatively) good bounds on coefficients in a geometric manner. We will see that the relationship is somewhat more complicated than $G(\beta)$ representing purely the excluded volume effects of the interaction, but nevertheless, it provides a useful starting point to understanding the convergence of such an expansion.

The second virial coefficient is $-\frac{1}{2}\beta_1$, where $\beta_1 = \int_0^\infty f(r)4\pi r^2 \, \mathrm{d}r$. If we cast it in the form:

$$\beta_1 = 4\pi \int_0^\infty r^2(e^{-\beta U(r)} - 1)\mathrm{d}r \tag{II.1.8}$$

We see that the linear relationship in temperature as $\beta a - b$ would require that a first order expansion of the exponential in the integral suffices for some region i.e. that $|U(r)|$ is small ($|U(r)| << \frac{1}{\beta}$) and for a temperature independent term for a region where $U(r)$ is large and positive ($U(r) >> \frac{1}{\beta}$). The integral over values in between these two cases should be minimal, so that it can be neglected.

In the first case, we would estimate

$$e^{-\beta U(r)} \approx 1 - \beta U(r) \tag{II.1.9}$$

This gives (in three dimensions):

$$a \approx 2\pi \int_{R_1} r^2 U(r)\mathrm{d}r \tag{II.1.10}$$

where $R_1$ is the region of validity of the first case. In the second case, the approximation is:

$$e^{-\beta U(r)} \approx 0 \tag{II.1.11}$$

This then gives:

$$b \approx 2\pi \int_{R_2} r^2 \mathrm{d}r \tag{II.1.12}$$

for $R_2$ the region of validity of the second case.

### II.1.1 Issues with the van der Waals Equation and the Law of Corresponding States

If we multiply both sides of the van der Waals Equation (II.1.2) by $\frac{V^2}{N^3}$, we get the expression:

$$\left(\frac{V^2}{N^2}P + a\right)\left(\frac{V}{N} - b\right) = \frac{V^2}{N^2}kT \tag{II.1.13}$$

For $v = \frac{V}{N}$ the reduced volume, we have the polynomial:

$$(v^2P + a)(v - b) = v^2kT \tag{II.1.14}$$

We divide through by $P$

$$\left(v^2 + \frac{a}{P}\right)(v - b) = v^2\frac{kT}{P} \tag{II.1.15}$$

and obtain the cubic equation in volume:

$$v^3 - \left(\frac{kT}{P} + b\right)v^2 + \frac{a}{P}v - \frac{ab}{P} = 0 \tag{II.1.16}$$

we can rearrange this to get an expression for $P$:

$$P = \frac{kTv^2 - av + ab}{v^3 - bv^2} \tag{II.1.17}$$

We see that at infinite reduced volume ($v = \infty$), $P = 0$ and the value of $P$ increases as $v$ decreases down to $b$, when $P = \infty$. If we calculate the reduced volume derivative, at fixed temperature $T$, we get:

$$\left(\frac{\partial P}{\partial v}\right)_T = -\frac{kTv^4 - 2v^3 + 4abv^2 + 2ab^2v}{(v^3 - bv^2)^2} \tag{II.1.18}$$

For large temperature, this is asymptotic to:

$$\frac{-kTv^4}{(v^3 - bv^2)^2} \tag{II.1.19}$$

which is negative and so for high enough temperatures pressure is monotonic in reduced volume. However, the cubic polynomial, may have turning points for lower temperature. The critical point of this behaviour is when the maxima and minima are at the same $v_c$, which happens at the critical temperature $T_c$ for strictly decreasing pressure, and the corresponding pressure is $P_c$. At such a point our equation

(II.1.16) is cast in the form:

$$(v - v_c)^3 = 0$$
$$v^3 - 3v_c v^2 + 3v_c^2 v - v_c^3 = 0 \tag{II.1.20}$$

Comparing the coefficients from (II.1.16), we get:

$$3v_c = \frac{kT_c}{P_c} + b \tag{II.1.21}$$
$$3v_c^2 = \frac{a}{P_c} \tag{II.1.22}$$
$$v_c^3 = \frac{ab}{P_c} \tag{II.1.23}$$

taking (II.1.23) / (II.1.22), we get:

$$b = \frac{v_c}{3} \tag{II.1.24}$$

substituting this into (II.1.21), we get:

$$\frac{8}{3} v_c = \frac{kT_c}{P_c} \tag{II.1.25}$$

We note that the experimental value of $\frac{v_c P_c}{kT_c}$ is around 0.3 and the value obtained using the van der Waals Equation of State is $\frac{3}{8}$, which are rather close. If we substitute for $a$ and $b$ from equations (II.1.22) and (II.1.24), respectively into equation (II.1.16), we obtain:

$$\left(v^2 P + 3P_c v_c^2\right)\left(v - \frac{v_c}{3}\right) = v^2 kT \tag{II.1.26}$$

we manipulate this equation using the relationship (II.1.25) to get:

$$\left(\left(\frac{v}{v_c}\right)^2 \left(\frac{P}{P_c}\right) + 3\right)\left(\frac{v}{v_c} - \frac{1}{3}\right) = \left(\frac{v}{v_c}\right)^2 \frac{kT}{v_c P_c}$$
$$\frac{P}{P_c} = \frac{8\frac{T}{T_c}}{3\frac{v}{v_c} - 1} - \frac{3}{\left(\frac{v}{v_c}\right)^2} \tag{II.1.27}$$

Some plots of isotherms for various values of $\frac{T}{T_c}$ are displayed in Figures II.1 and II.2

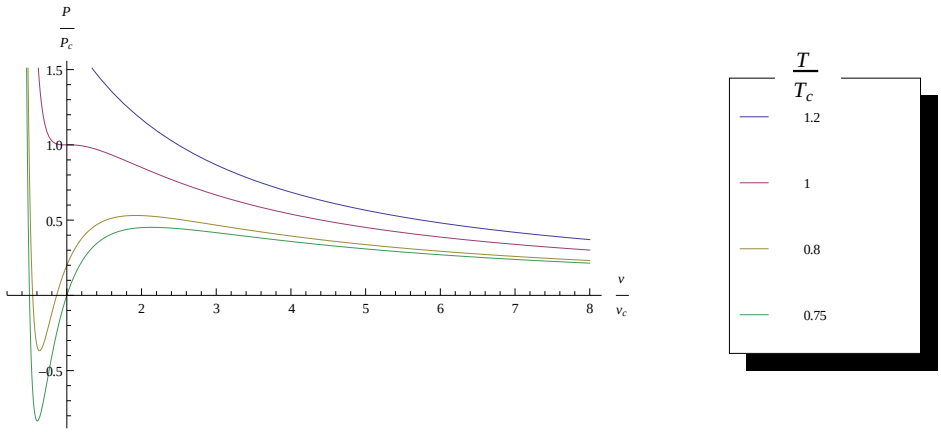



Figure II.1: Plots for isotherms of equation (II.1.27)

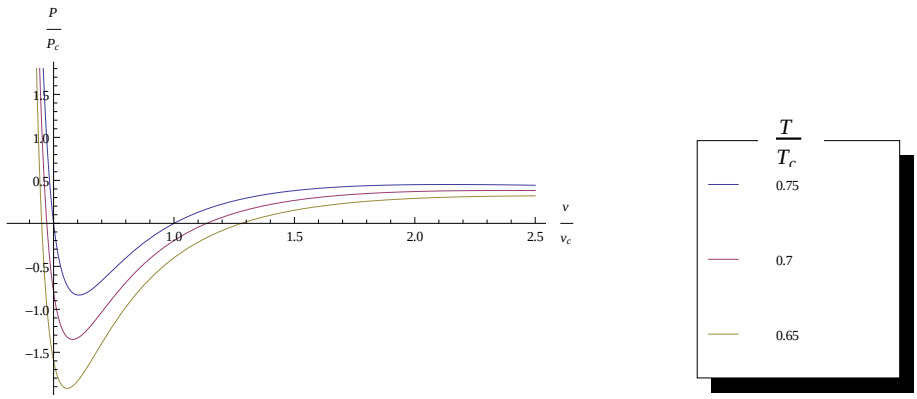



Figure II.2: Plots for isotherms of equation (II.1.27)

We can see very clearly in Figure II.2 that we have regions where the pressure increases with increasing reduced volume. This makes very little physical sense and it is where Maxwell's idea of altering such graphs to add a straight line between two points of equal value of $\frac{P}{P_c}$, where the area below and above are equal to mark the effects of a liquid-gas phase co-existence. The tricritical point of the system is evident in Figure II.1, where we can clearly see the point of inflection at $P = P_c$ and $v = v_c$.

The paper [LePen66] makes rigorous and generalises the derivation of a van der Waals equation of state given by van Kampen in [Kam64]. This uses cell approximations and Kac potentials. There are also exact results given on Lennard-Jones potentials in the papers [Jon24, LeDe37].

## II.2 Mayer's Second Theorem

This section is based on derivations attributed to Groeneveld and Mayer in the book by McCoy [McCoy10] and the work of Mayer and Mayer [MMay40]. The comments made throughout have been informed by the work of Pulvirenti and Tsagkarogiannis [PT12] and the paper by Vasil'ev and Radzhabov [VaRa75]. Connections to combinatorics are informed by the work of Leroux [Ler04] and Faris [Far10, Far11].

In Chapter I, the relationship between connected graphs and the cluster expansion was derived and a similar relationship can be derived for the virial coefficients, but the type of graphs in this case is two connected graphs. Two-connected graphs are also called two-vertex connected graphs and star graphs in the literature. These are distinct from one particle irreducible graphs, which are two-edge connected graphs. The main idea is that when we are calculating the terms in the cluster expansion and we take the infinite volume limit, the integral over one of the variables becomes redundant, due to the translational symmetry of the potential. This provides us with a way of factorising the integrals we have in terms of those which are termed 'irreducible'.

The motivating example is one of the integrals we get in the third cluster coefficient:

$$\int_{\mathbb{R}^3} \mathrm{d}q_1 \int_{\mathbb{R}^3} \mathrm{d}q_2 \int_{\mathbb{R}^3} \mathrm{d}q_3 \, f_{1,2} f_{1,3} = \int_{\mathbb{R}^3} \mathrm{d}q_1 \left( \int_{\mathbb{R}^3} f_{1,2} \mathrm{d}q_2 \right) \left( \int_{\mathbb{R}^3} f_{1,3} \mathrm{d}q_3 \right) \tag{II.2.1}$$

We change our variables in the appropriate way -since the $f_{1,2}$ is assumed to depend only on $|q_1 - q_2|$ - to get an independent integral, which will give us the volume term, and two factorised integrals for $\gamma = |q_1 - q_2|$ and $\mu = |q_1 - q_3|$. If we define:

$$\beta_1 := 2\pi \int_0^\infty f(q) \mathrm{d}q \tag{II.2.2}$$

then this term contributes $\beta_1^2$, after dividing by the volume term.

If we instead consider the integral:

$$\beta_2 := \int_{\mathbb{R}^3} \mathrm{d}q_1 \int_{\mathbb{R}^3} \mathrm{d}q_2 \int_{\mathbb{R}^3} \mathrm{d}q_3 \, (f_{1,2} f_{1,3} f_{2,3}) \tag{II.2.3}$$

we see that such a factorisation is not possible, since there is no splitting of the $f$-functions into two where they have only one point in common. Such an integral is called *irreducible*.

We denote the irreducible integral of order $i+1$ as $\beta_i$. We can then write our

cluster integrals in terms of such factors. The remarkable consequence of introducing such integrals is that they provide precisely the coefficients of the virial expansion.

**Theorem II.2.1** (Mayer's Second Theorem)**.** *In the infinite volume limit of the classical gas model described above, we have the following virial expansion:*

$$\beta P(\rho) = \rho - \sum_{k=1}^{\infty} \frac{k}{(k+1)!} \beta_k \rho^{k+1} \tag{II.2.4}$$

*where the $\beta_k$ are the irreducible integrals of order $k+1$.*

In order to prove this theorem, we need to understand how to obtain these factorisations in terms of properties of the graphs we use to represent the functions. The main idea is that it is precisely what are known as 'articulation points' about which we can take such a factorisation.

**Definition.** *An articulation point on a connected graph $G$ is a vertex such that its removal (and the removal of all adjacent edges) renders the graph disconnected.*

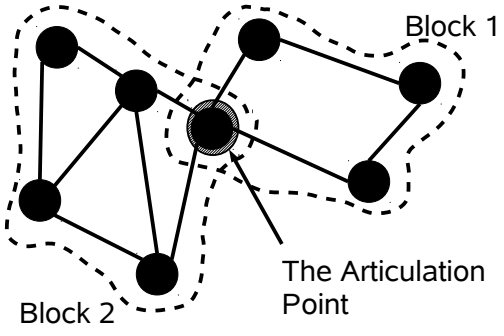



Figure II.3: A simple graph with one articulation point

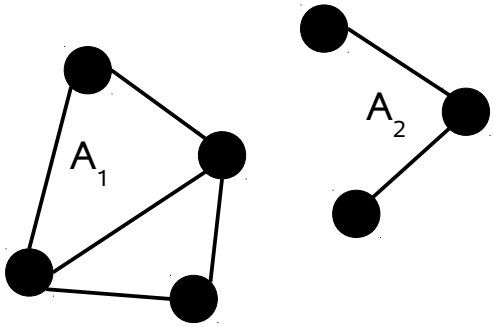



Figure II.4: How we separate the remaining vertices

A helpful interpretation of articulation points in this section is to think of them as the vertex at which we can attach blocks. In Figure II.3, we see that we have two smaller connected graphs which are 'glued' at the marked articulation point. This is related to the factorisation property, since the only points we can make such a factorisation of integrals of products of the $f_{i,j}$ is at articulation points.

When an articulation point, $a$, is removed our vertex set is naturally split into (at least) two parts (corresponding to connected subgraphs). Let us call one

individual part $A_1$ and the remaining vertices $A_2$. If we write $B_1 = A_1 \cup \{a\}$ and $B_2 = A_2 \cup \{a\}$ and we notice that all edges in our original graph $g$ have endpoints either both in $B_1$ or both in $B_2$ and no edges with one endpoint in $A_1$ and the other in $A_2$. This is shown in an example in Figure II.4.

If we write

$$f_{g,S} := \prod_{e \in E(g) \cap S^{(2)}} f_e \tag{II.2.5}$$

We can then write:

$$\prod_{i \in [n]} \left( \int_{\mathbb{R}^3} \mathrm{d}q_i \right) f_{g,[n]} = \int_{\mathbb{R}^3} \mathrm{d}q_a \prod_{i \in A_1} \left( \int_{\mathbb{R}^3} \mathrm{d}q_i \right) f_{g,B_1} \prod_{i \in A_2} \left( \int_{\mathbb{R}^3} \mathrm{d}q_i \right) f_{g,B_2} \tag{II.2.6}$$

We thus see that this integral is factorised into the product of the two integrals, since they are both independent of the single coordinate $a$. Of course this can be repeated at all articulation points and if we have multiple sets $A_i$ from the splitting of the graph into connected components upon removal of $a$, then we have the integral factorising into that many products. Furthermore, it is only at these articulation points we can make such a factorisation, or else we will have edges between the components and the integral would not factorise. This, thus, motivates the investigation of articulation points and two-connected graphs (those without articulation points) in order to understand the virial expansion.

Mayer's Second Theorem is approached in three stages, where various weighted graph functions are introduced and their functional relations to each other are conveyed, in order to have the interpretation of the virial coefficients.

### II.2.1 Part One: D-graphs

We first define the Ursell functions, which represent the cluster coefficients:

**Definition** (Ursell Functions)**.** *The Ursell function of $k$ positions $U_k(1, \cdots, k)$, maps $U : \mathbb{R}^{3k} \to \mathbb{R}$ and is defined as:*

$$U_k(1, \cdots .k) := \sum_{g \in \mathcal{C}[k]} \prod_{e \in E(g)} f_e \tag{II.2.7}$$

We want to understand the connection between the irreducible diagrams and the cluster diagrams we have defined above. The first step in doing so, is to consider graphs which have some relation to both. The type of graph to consider is the $D$-graph defined below:

**Definition** ($D$-graphs)**.** *Define $\mathcal{D}[k]$ as the collection of connected graphs where* 1

*is not an articulation point. The related function* $D_k(1,\cdots,k)$, $D:\mathbb{R}^{3k}\to\mathbb{R}$ *is defined by:*

$$D_k(1,\cdots,k) := \sum_{g\in\mathcal{D}[k]}\prod_{e\in E(g)} f_e \tag{II.2.8}$$

A graph in $\mathcal{D}[k]$ can be interpreted as having 1 as a special 'ghost' vertex, which is internal to a two-connected component inside the overall graph. Each vertex in the two-connected component, may itself be an articulation point, with its own connected graph emanating from this point, or it may just be another internal vertex. In either case each of the connected graphs (taking the connected graph of one vertex, if it is an internal vertex) can be viewed as one rooted at the point in the block. Two-connected components are called 'blocks' and are maximal subgraphs of the main graph which are two-connected, i.e. without an articulation point. The situation of D-graphs is illustrated in Figure II.5. The combinatorial interpretation of this is as a $\mathcal{B}'(\mathcal{C}^\bullet)$-structure, which is explained in Chapter V.

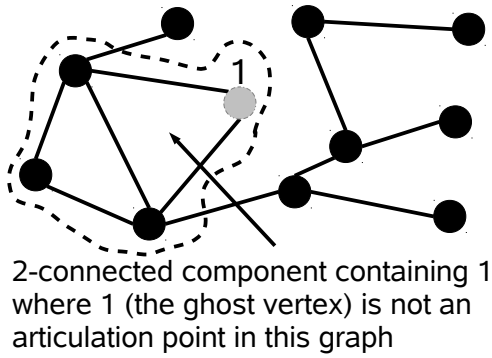



Figure II.5: An Example of a D-graph

We rewrite the sum over all connected graphs on $l$ vertices, according to how 1 can be an articulation point. The recursion involves the natural splitting of the vertices into two sets, upon the removal of the vertex labelled 1 in the graph. The vertices that remain connected to 2 and thus each other when 1 is removed, are understood to be in the $\mathcal{D}$-graph. The disjoint set of points that are separated from those in the $\mathcal{D}$-graph, with the removal of the vertex labelled 1 are included in a separate connected graph involving the vertex 1. This can be expressed in terms of

the graph functions:

$$U_l(1,\cdots,l) = D_{l-1}(1;2,\cdots,l)U_1(1) + \sum_{3\leq j\leq l} D_{l-2}(1,2,\cdots,\hat{j},\cdots,l)U_2(1,j) + \sum_{3\leq j_1<j_2\leq l} T_{l-3}(1;,2,\cdots,\hat{j}_1,\cdots,\hat{j}_2,\cdots l)U_3(1,j_1,j_2)+\cdots + D_1(1;2)U_l(1,3,\cdots,l) \tag{II.2.9}$$

The term involving $U_k$ has $k-1$ variables chosen from $\{3,\cdots,l\}$ and hence there are $\binom{l-2}{k-1}$ such terms.

We define the integrated quantity:

$$d_l = \frac{1}{l!}\int D_l(1;2,\cdots,l+1)\mathrm{d}q_2\cdots\mathrm{d}q_{l+1} \tag{II.2.10}$$

where the integrals are over the whole space.

**Remark 14.** *We note that it is necessary that the integrals are over the whole space and not just a finite region, in order to obtain the factorisation property, explained in the beginning of this section. Indeed, from Canonical Ensemble calculations in the paper by Pulvirenti and Tsagkarogiannis [PT12], they find that there are extra terms present in the virial expansion when we have only finite volume forms. We must integrate over the whole space so the exact position of the point* 1 *is irrelevant. In the paper, for finite volume, they need to make the assumption of translation invariance of the potential. It is also important to remark that this is if we desire the the relationship with two-connected graphs to be expressed exactly in the way presented here.*

We use the definitions of $b_k$ and integrate both sides of (II.2.9), noting that we have the factorisation property over 1, since it is an articulation point in each case, except the first term, for which there is no splitting. We thus obtain the expression:

$$l!b_l = \sum_{k=1}^{l-1} k!b_k(l-k)!d_{l-k}\binom{l-2}{k-1} \tag{II.2.11}$$

hence

$$l(l-1)b_l = \sum_{k=1}^{l-1} kb_k(l-k)d_{l-k} \tag{II.2.12}$$

multiply both sides by $z^{l-1}$ and sum from two to infinity.

Notice for the LHS

$$\sum_{l=1}^{\infty} z^{l-1} l(l-1) b_l = z \frac{\partial}{\partial z} \frac{1}{z} \sum_{l=1}^{\infty} l z^l b_l = z \frac{\partial}{\partial z} \frac{\rho(z)}{z} \tag{II.2.13}$$

and we interchange the orders of summation on the RHS and introduce the generating function:

$$D(z) = \sum_{n=1}^{\infty} d_n z^n \tag{II.2.14}$$

to find that

$$\begin{aligned} \sum_{l=1}^{\infty} z^{l-1} \sum_{k=1}^{l-1} k b_k (l-k) d_{l-k} &= \sum_{k=1}^{\infty} k b_k z^k \sum_{l-k=1}^{\infty} z^{l-k-1} (l-k) d_{l-k} \\ &= \rho(z) \frac{\partial}{\partial z} D(z) \end{aligned} \tag{II.2.15}$$

putting this together we get:

$$z \frac{\partial}{\partial z} \left( \frac{\rho(z)}{z} \right) = \rho(z) \frac{\partial}{\partial z} D(z) \tag{II.2.16}$$

This can be solved, using initial conditions $D(0) = 0$ and $\lim_{z \to 0} \frac{\rho(z)}{z} = 1$, we find that:

$$\rho(z) = z e^{D(z)} \tag{II.2.17}$$

**Remark 15.** *In chapter V, we derive the combinatorial relation:*

$$\mathcal{C}^{\bullet} = X \star \mathcal{E} \circ \mathcal{B}'(\mathcal{C}^{\bullet}) \tag{II.2.18}$$

*which corresponds precisely to this explicit relationship between the density and the D-function. Since $\mathcal{C}^{\bullet}$ corresponds to density and as remarked earlier $\mathcal{B}'(\mathcal{C}^{\bullet})$ corresponds to $\mathcal{D}$ graphs. The composition with $\mathcal{E}$ corresponds to taking an exponential and is the idea of partitioning a set into 'substructures' and the extra $X$ is the 'rooted point'.*

### II.2.2 Part Two: Y-graphs

**Definition** (The Graphs $Y_l^{(s)}(\{\alpha\}_s; 1, 2, \cdots, l)$)**.** *We define the function $Y_l^{(s)}(\{\alpha\}_s; 1, 2, \cdots, l)$ as the sum over all Mayer weights for graphs on $l+s$ points with the following properties:*

*i) No path connects any pair of points in the set $\{\alpha\}_s$*

*ii) the diagram becomes connected if all points $\{\alpha\}$ are connected by lines.*

The $Y$-graphs are related to the correlation functions and the recursion relations we gain here should be compared with the general integral equations for correlation functions in Section II.3. There are subtle differences as these are constructed only in the sense of being products of connected graphs with a corresponding 'ghost vertex' in the set represented by $\alpha$. They are somewhat simpler than correlation functions but nevertheless the way in which they are treated is quite similar due to this collection of 'ghost vertices'. The points inside of the correlation function are related to the points which are declared ghost vertices in $\{\alpha\}$.

There is a recursion relation which expresses the graphs $Y_l^{(s+1)}$ in terms of graphs in $Y_{l-k+1}^{(s)}$ and $U_k$ for $k = 1, \cdots, l+1$. This is obtained by adding the point $1'$ to the set of $s$ points in $\{\alpha\}$, where this new added point is connected to $k-1$ of the points $1, 2, \cdots, l$. Therefore:

$$\begin{aligned} Y_l^{(s+1)}(\{1', \{\alpha\}_s\}_{s+1}; 1, 2, \cdots, l) = {} & Y_l^{(s)}(\{\alpha\}_s; 1, \cdots, l) U_1(1') \\ & + \sum_{1 \le j \le l} Y_{l-1}^{(s)}(\{\alpha\}_s; 1, \cdots, \hat{j}, \cdots, l) U_2(1', j) \\ & + \sum_{1 \le j_1 < j_2 \le l} Y_{l-2}^{(s)}(\{\alpha\}_s; 1, \cdots, \hat{j}_1, \cdots, \hat{j}_2, \cdots, l) U_2(1', j_1, j_2) \\ & + \cdots + Y_0^{(s)}(\{\alpha\}_s;) U_{l+1}(1', 1, \cdots, l) \end{aligned} \tag{II.2.19}$$

The sum including $U_k$ has $k-1$ points freely chosen from $l$ possible points and so the number of terms in the sum is: $\binom{l}{k-1}$.

We now define:

$$y_l^{(s)} = \frac{1}{l!} \int Y_l^{(s)}(\{\alpha\}_s; 1, \cdots, l) \, \mathrm{d}^D q_1 \cdots \mathrm{d}^D q_l \tag{II.2.20}$$

The integrals are over the whole space and not a finite volume $V$. We integrate our recursion over the whole space. The functions $Y_{l-k}^{(s)}$ and $U_{k+1}$ have no points in common and so the integral of the product factorises into the product of the integrals. Using the definitions of $y_l^{(s)}$ and $b_l$, we find:

$$l! y_l^{(s+1)} = \sum_{k=1}^{l+1} k! b_k (l-k+1)! y_{l-k+1}^{(s)} \frac{l!}{(k-1)!(l-k+1)!} \tag{II.2.21}$$

hence:

$$y_l^{(s+1)} = \sum_{k=1}^{l+1} k b_k y_{l-k+1}^{(s)} \tag{II.2.22}$$

We define the generating function:

$$Y^{(s)}(z) = \sum_{l=0}^{\infty} y_l^{(s)} z^l \tag{II.2.23}$$

We multiply the left hand side of (II.2.22) by $z^l$ and sum over $l$, to obtain:

$$Y^{(s+1)}(z) = \rho(z) \frac{Y^{(s)}(z)}{z} \tag{II.2.24}$$

We also note that:

$$Y_l^{(l)}(\{1'\}_1; 1, \cdots, l) = U_{l+1}(1', 1 \cdots, l) \tag{II.2.25}$$

Integrating over all coordinates, we get:

$$y_l^{(1)} = (l+1) b_{l+1} \tag{II.2.26}$$

and hence we get the relationship between the two generating functions as:

$$Y^{(1)}(z) = \frac{\rho(z)}{z} \tag{II.2.27}$$

We solve the recursion relation by induction to get:

$$Y^{(s)}(z) = \left( \frac{\rho(z)}{z} \right)^s \tag{II.2.28}$$

**Remark 16.** *In terms of combinatorial species, covered in Chapter V we can interpret the $Y^{(s)}$-structure as being the composition of $X^s$ - the species which for any set of size $s$ gives all permutations on the set as the structure, but for sets of any other size gives the empty set - with $\mathcal{C}'$ , which are connected graphs with a ghost vertex. We understand $\{\alpha\}_s$ as representing a set of ghost vertices, from which we can freely choose a 'ghost' for each subset of the graph we get by considering connected components of the final graph. Furthermore, the relationship in the recursion can be expressed as: $X^{s+1} \circ \mathcal{C}' = \mathcal{C}' \star X^s \circ \mathcal{C}'$, which is somewhat trivial. We identify the operation of composition by $X^s$ as multiplying a species by itself $s$ times. The species $\mathcal{C}'$ is readily identified as relating to the expression $\frac{\rho(z)}{z}$ to give the result in a transparent way.*

### II.2.3 Part Three: Irreducible Diagrams

The term irreducible diagram is synonymous with a two-connected graph. Of course, now we have developed our knowledge of graphs which relate to splitting connected graphs in various ways, we would now like to understand and see where the irreducible diagrams come in.

**Definition** (The Irreducible Diagrams (Husimi Functions): $V_l(1,\cdots,l)$)**.** *Two connected graphs, which are indicated by $\mathcal{B}$, have a corresponding function formed from the Mayer weights. These functions are called Husimi functions.*

*The Husimi function, denoted $V_l(1,\cdots,l)$, is a function $V:\mathbb{R}^{3l}\to\mathbb{R}$ defined by:*

$$V_l(1,\cdots,l) := \sum_{g\in\mathcal{B}[l]}\prod_{e\in E(g)} f_e \tag{II.2.29}$$

We construct a recursion relation between $D_l(1';1,\cdots,l)$ in terms of $V_{l+1-k}$ and $Y_k^{(l-k)}$ for $k=0,\cdots,l-1$. We do this by considering $1'$ as a special point and dividing the graphs into the graphs that are two-connected to $1'$ and the rest, where the points which are two-connected to $1'$, are connected to the rest of the points through at least one articulation point. We count in terms of the number of such points that are connected to $1'$ through articulation points. If we recall Figure II.5, we see that each point in the two-connected component (except $1'$), can be understood as belonging to the set $\alpha$ for $Y$, since we have separate connected graphs emanating from each point. The internal structure is obviously that of an irreducible graph, by definition. Thus we find:

$$\begin{aligned}
D_l(1';1,\cdots,l) &= V_{l+1}(1',1,\cdots,l)Y_0^{(l)}(\{1,\cdots,l\}_l,)\\
&+\sum_{1\le j\le l} V_l(1',1,\cdots,\hat{j},\cdots,l)Y_1^{(l-1)}(\{1,\cdots,\hat{j},\cdots,l\}_{l-1};j)\\
&+\sum_{1\le j_1<j_2\le l} V_{l-1}(1',1,\cdots,\hat{j}_1,\cdots,\hat{j}_2,\cdots,l)Y_2^{(l-2)}(\{1,\cdots,\hat{j}_1,\cdots,\hat{j}_2,\cdots,l\}_{l-2};j_1,j_2)\\
&+\cdots+\sum_{1\le j\le l} V_2(1',j)Y_{l-1}^{(1)}(\{j\}_1;1,\cdots,\hat{j},\cdots,l)
\end{aligned} \tag{II.2.30}$$

For the term in the sum containing $Y_{l-j}^{(j)}$ the $j$ points in the set $\alpha$ are freely chosen from $l$ possible points and so contributes $\binom{l}{j}$ terms. We now integrate over $q_1,\cdots,q_l$ in (II.2.30) with $q_{1'}$ fixed. We use the fact that the integral of the product

$V_{l-k}Y_{k+1}^{(l-k-1)}$ factorises and is independent of $q_{1'}$, to find:

$$l!d_l = \sum_{j=1}^{l} \beta_j (l-j)! y_{l-j}^{(j)} \frac{l!}{j!(l-j)!} \tag{II.2.31}$$

and thus:

$$d_l = \sum_{j=1}^{l} \frac{\beta_j}{j!} y_{l-j}^{(j)} \tag{II.2.32}$$

Then multiply (II.2.32) by $z^l$ and sum from one to infinity to get:

$$D(z) = \sum_{j=1}^{\infty} z^j \frac{\beta_j}{j!} Y^{(j)}(z) \tag{II.2.33}$$

We use our expression for $Y^{(j)}(z)$ from (II.2.28) to get the expression:

$$D(z) = \sum_{j=1}^{\infty} \rho^j \frac{\beta_j}{j!} \tag{II.2.34}$$

We then use the result (II.2.17) ( $\rho \exp(-D(z)) = z$) to obtain the expansion:

$$z(\rho) = \rho \exp\left( - \sum_{j=1}^{\infty} \frac{\beta_j}{j!} \rho^j \right) \tag{II.2.35}$$

We use this inversion formula to obtain an expression for $p$ to give a virial expansion. We know that:

$$\frac{\partial}{\partial z}(\beta p(z)) = \frac{\rho(z)}{z} \tag{II.2.36}$$

and hence we have:

$$\frac{\partial}{\partial \rho}(\beta p(\rho)) = \frac{\partial z}{\partial \rho} \frac{\partial}{\partial z}(\beta p(z)) = \frac{\partial z}{\partial \rho} \frac{\rho}{z} = \frac{\partial \ln z}{\partial \ln \rho} \tag{II.2.37}$$

We then use (II.2.35) to obtain an expression for $\ln z$ in terms of $\rho$:

$$\ln z = \ln \rho - \sum_{k=1}^{\infty} \frac{\beta_k}{k!} \rho^k \tag{II.2.38}$$

Hence we find:

$$\frac{\partial \ln z}{\partial \ln \rho} = 1 - \frac{\partial \rho}{\partial \ln \rho} \frac{\partial}{\partial \rho} \left( \sum_{k=1}^{\infty} \frac{\beta_k}{k!} \rho^k \right) = 1 - \sum_{k=1}^{\infty} k \frac{\beta_k}{k!} \rho^k \tag{II.2.39}$$

We integrate this expression to get the virial expansion:

$$\beta p(\rho) = \rho - \sum_{k=1}^{\infty} \frac{k}{(k+1)!} \beta_k \rho^{k+1} \tag{II.2.40}$$

We thus have that the virial coefficients are represented by the irreducible integrals we found at the beginning of this section. Although the path to such a solution has been somewhat convoluted. In Chapter V, a more direct route of understanding the different coefficients is presented, through the language of combinatorial species of structure.

The method relies firstly on understanding a relationship between density expansions and the expansion of $D$-graphs. This follows from considering rooted connected graphs as being made from the different possibilities of graphs, where we split the graph into two-connected components from the root vertex 1. We then understand how the individual $D$-graphs and be split into the two-connected graph containing the root 1 and the other connected graphs emanating from the articulation points of this privileged two-connected graph in the original graph. This gives us our product of Husimi functions and the function corresponding to $Y$-graphs. The rest relies on understanding where particular coefficients and definitions of the thermodynamic functions come into play.

## II.3 The Approach of Kirkwood and Salsburg and other Integral Equations

The approach shared by Ruelle in [Rue69] and Gruber and Kunz in [GK71] in obtaining an estimate on correlation functions in order to understand the convergence of the virial expansion is through Kirkwood Salsburg equations, which were first introduced in [KiSa53]. Ruelle obtains important bounds on the correlation functions from these integral equations and uses these to understand the virial expansion through a Cauchy integral representation of the coefficients. There is also one other important set of integral equations, which is used to obtain bounds on correlation functions. these are called the Mayer Montroll equations and can be found in [MMon41]. Furthermore, Lebowitz and Percus derived general integral equations, which interpolate between these two sets of equations in [LePer63].

This section presents a simple bound on the radius of convergence for the virial expansion, obtained by Ruelle [Rue69], in order to motivate the approach of Chapter IV.

Assume the potential $\mathcal{U}$ can be expressed as a sum over pair potentials $\Phi$.

$$\mathcal{U}(\mathbf{x})_n = \sum_{1\leq i<j\leq n} \Phi(\mathbf{x}_i - \mathbf{x}_j) \tag{II.3.1}$$

$$\psi(\mathbf{x})_n = \exp\left(-\beta \mathcal{U}(\mathbf{x})_n\right) \tag{II.3.2}$$

Using this notation, we write the grand canonical partition function as:

$$\Xi_\Lambda(z,\beta) = \sum_{n=0}^{\infty} \frac{z^n}{n!} \int_{\Lambda^n} \mathrm{d}\mathbf{x}_1 \cdots \mathrm{d}\mathbf{x}_n \, \psi(\mathbf{x})_n \tag{II.3.3}$$

**Definition.** *We define the* m-point correlation function $\rho_\Lambda(\mathbf{x})_m$ *as the probability density (with respect to Lebesgue measure) of finding* $m$ *different particles at locations* $\mathbf{x}_1 \cdots \mathbf{x}_m \in \Lambda$. *We have:*

$$\rho_\Lambda(\mathbf{x})_m = \Xi_\Lambda(z,\beta)^{-1} \sum_{n=1}^{\infty} \frac{z^{m+n}}{n!} \int_{\Lambda^n} \mathrm{d}\mathbf{x}_{m+1} \cdots \mathrm{d}\mathbf{x}_{m+n} \, \psi(\mathbf{x})_{m+n} \tag{II.3.4}$$

*We denote by* $\rho(\mathbf{x})_m$ *the thermodynamic limit (as* $\Lambda \to \mathbb{R}^D$ *in the sense of van Hove) of* $\rho_\Lambda(\mathbf{x})_m$.

We require the stability condition. That is: there exists $B \geq 0$, such that $\forall n$ and $\mathbf{x}_1 \cdots \mathbf{x}_n \in \mathbb{R}^D$

$$\mathcal{U}(\mathbf{x})_n = \sum_{1\leq i<j\leq n} \Phi(\mathbf{x}_i - \mathbf{x}_j) \geq -Bn \tag{II.3.5}$$

We also assume that our potential is *regular* or *tempered*, that is:

$$C(\beta) = \int |e^{-\beta\Phi(\mathbf{x})} - 1| \mathrm{d}^D\mathbf{x} < +\infty \tag{II.3.6}$$

for some $\beta > 0$ and therefore for all $\beta > 0$.

### II.3.1 The Kirkwood Salsburg Equations

The Kirkwood Salsburg equations are introduced in the context of the subset gas in Section I.3. They arise from the consideration of correlation functions with a different number of particles and how we may 'integrate out' some of the particles to obtain a relation. The algebra of such equations is general and is used by Ruelle [Rue69] and Poghosyan and Ueltschi [PoUe09] in the context of a classical gas in particular. This section indicates the specialisation of the general Kirkwood Salsburg equations to the case of a classical gas.

**Remark 17** (The Algebraic Approach of Ruelle)**.** *In this section, we use the algebraic approach of Ruelle [Rue69], where* $\boldsymbol{\phi} = (\phi(x)_n)_{n\in\mathbb{N}}$ *is a sequence of complex valued functions. The arguments of the functions are now sequences in* $\mathbb{R}^3$*. This is to fit the classical gas case that is considered here.*

For a product of characteristic functions, we write $\chi_\Lambda(\mathbf{x})_n = \prod_{i=1}^{n} \chi_\Lambda(\mathbf{x}_i)$. The $m$-point correlation function may be rewritten in the convenient form:

$$\rho_\Lambda(\mathbf{x})_m = \Xi_\Lambda(z,\beta)^{-1} \sum_{n=0}^{\infty} \frac{z^{m+n}}{n!} \int \mathrm{d}^D(\mathbf{x})_{[m+1,m+n]}\, \chi_\Lambda(\mathbf{x})_{m+n} \psi(\mathbf{x})_{m+n} \tag{II.3.7}$$

We use the notation for interactions involving particle $j$:

$$W^j(\mathbf{x})_m = \sum_{i\neq j} \Phi(\mathbf{x}_i - \mathbf{x}_j) \tag{II.3.8}$$

$$\Psi^j(\mathbf{x})_m = \exp\left(-\beta W^j(\mathbf{x})_m\right) \tag{II.3.9}$$

and $(\mathbf{x})'_{m-1} = (\mathbf{x}_2, \cdots, \mathbf{x}_m)$.

We define the (Kirkwood Salsburg) kernel:

$$K(\mathbf{x}_1, (\mathbf{y})_n) = \prod_{j=1}^{n} \left(\exp\left(-\beta\Phi(\mathbf{y}_j - \mathbf{x}_1)\right) - 1\right) \tag{II.3.10}$$

So that the Kirkwood Salsburg equations can be written in the form:

$$\rho_\Lambda(\mathbf{x}_1) = \chi_\Lambda(\mathbf{x}_1) z \left(1 + \sum_{n=1}^{\infty} \frac{1}{n!} \int \mathrm{d}^D(\mathbf{y})_n\, K(\mathbf{x}_1, (\mathbf{y})_n) \rho_\Lambda(\mathbf{y})_n\right) \tag{II.3.11}$$

$$\rho_\Lambda(\mathbf{x})_m = \chi_\Lambda(\mathbf{x})_m z \Psi^1(\mathbf{x})_m \left(\rho_\Lambda(\mathbf{x})'_{m-1} + \sum_{n=1}^{\infty} \frac{1}{n!} \int \mathrm{d}^D(\mathbf{y})_n\, K(\mathbf{x}_1, (\mathbf{y})_n) \rho_\Lambda((\mathbf{x})'_{m-1}, (\mathbf{y})_n)\right) \tag{II.3.12}$$

### II.3.2 Banach Spaces

We wish to understand these inductive relations as operations in a Banach space so that we can understand fixed point theorems and obtain an estimate on the norms of the Kirkwood Salsburg operator. We define the Banach Space $B_\xi$ with the norm:

$$\|\boldsymbol{\phi}\|_\xi = \sup_{n\geq 1} \left(\xi^{-n} \sup_{(\mathbf{x})_n \in \mathbb{R}^{nD}} |\phi(\mathbf{x})_n|\right) \tag{II.3.13}$$

The Banach Space is the set of sequences for which this norm is finite. From stability and the equations for $\rho_\Lambda(x)_n$ (II.3.7) we have (by bounding $\psi(\mathbf{x})_{m+n} \leq e^{\beta B(m+n)}$ and evaluating the integral as the volume multiplied by this maximum):

$$0 \leq \rho_\Lambda(\mathbf{x})_n \leq \Xi_\Lambda(z,\beta)^{-1} \exp\left(zV(\Lambda)e^{\beta B}\right)(ze^{\beta B})^n \tag{II.3.14}$$

So $\boldsymbol{\rho}_\Lambda \in E_\xi$ if $\xi \geq ze^{\beta B}$. We then can rewrite the Kirkwood Salsburg equations as an equation for $\boldsymbol{\rho}_\Lambda \in E_\xi$.

We define the linear operator $\boldsymbol{\chi}_\Lambda$ on $E_\xi$ by:

$$\boldsymbol{\chi}_\Lambda \boldsymbol{\phi} := (\chi_\Lambda(x)_n \phi(x)_n)_{n\geq 1} \tag{II.3.15}$$

We then define $\mathbf{K}$ by:

$$(\mathbf{K}\boldsymbol{\phi})(x_1) = \sum_{n=1}^{\infty} \frac{1}{n!} \int \mathrm{d}(y)_n \, K(x_1,(y)_n)\phi(y)_n \tag{II.3.16}$$

$$(\mathbf{K}\boldsymbol{\phi})(x)_m = \Psi^1(x)_m \times \left(\phi(x)'_{m-1} + \sum_{n=1}^{\infty} \frac{1}{n!} \int \mathrm{d}(y)_n \, K(x_1,(y)_n)\phi((x)'_{m-1},(y)_n)\right) \tag{II.3.17}$$

We have $|(\mathbf{K}\boldsymbol{\phi})(x)_m| \leq \|\boldsymbol{\phi}\|_\xi \exp\left(\xi C(\beta)\right)(e^{2\beta B}\xi)^{m-1}$ and hence $\mathbf{K}$ maps $E_\xi$ into $E_{e^{2\beta B}\xi}$. We therefore can write the Kirkwood Salsburg equation as an equation in $E := \cup_{\xi>0} E_\xi$:

$$\boldsymbol{\rho}_\Lambda = z\boldsymbol{\chi}_\Lambda \boldsymbol{\alpha} + z\boldsymbol{\chi}_\Lambda \mathbf{K}\boldsymbol{\rho}_\Lambda \tag{II.3.18}$$

where $\boldsymbol{\alpha}$ is defined by:

$$\begin{aligned} \boldsymbol{\alpha}(x_1) &= 1 \\ \boldsymbol{\alpha}(x)_n &= 0 \text{ for } n \geq 2 \end{aligned} \tag{II.3.19}$$

We would like an equation to hold within a single $E_\xi$.
We know that $\sum_{i=1}^{m} W^i(x)_m = 2U(x)_m \geq -2mB$.
This implies that for each $(x)_m \in \mathbb{R}^{mD}$ we can choose a $j$ such that $W^j(x)_m \geq -2B$.
We construct the operator $\boldsymbol{\Pi}$ on $E_\xi$, which for each $(x)_m$, replaces $\phi(x)_m$ by $\phi(x_{i_1},\cdots x_{i_m})$, where the permutation $\pi : (1,\cdots,m) \to (i_1,\cdots,i_m)$ is chosen so that $W^{i_1}(x)_m \geq -2B$. Since $\rho_\Lambda(x)_m$ is symmetric in its arguments, we may use instead the equations:

$$\boldsymbol{\rho}_\Lambda = z\boldsymbol{\chi}_\Lambda \boldsymbol{\alpha} + z\boldsymbol{\chi}_\Lambda \boldsymbol{\Pi}\boldsymbol{K}\boldsymbol{\rho}_\Lambda \tag{II.3.20}$$

Then we have the inequality:

$$|(\mathbf{\Pi K \phi})(x)_m| \leq \|\boldsymbol{\phi}\|_\xi e^{2\beta B} \exp\left(\xi C(\beta)\right) \xi^{m-1} \tag{II.3.21}$$

so that $\mathbf{\Pi K}$ maps $E_\xi$ into itself with operator norm:

$$\|\mathbf{\Pi K}\|_\xi = \sup_{\|\boldsymbol{\phi}\|_\xi = 1} \|\mathbf{\Pi K \phi}\|_\xi \leq e^{2\beta B} \xi^{-1} \exp\left(\xi C(\beta)\right) \tag{II.3.22}$$

If we consider the equation $\boldsymbol{\rho} = z\boldsymbol{\alpha} + z\mathbf{\Pi K}\boldsymbol{\rho}$, then we see that the kernels of the equations have norms smaller than $|z|e^{2\beta B}\xi^{-1}\exp\left(\xi C(\beta)\right)$ and therefore equations (II.3.23) and (II.3.24) have a unique solution in $E_\xi$ as soon as $|z| < e^{-2\beta B - 1} C(\beta)^{-1}$, where we have taken $\xi = C(\beta)^{-1}$.

$$\boldsymbol{\rho}_\Lambda = \left(\mathbb{1} - z\boldsymbol{\chi}_\Lambda \mathbf{\Pi K}\right)^{-1} z\boldsymbol{\chi}_\Lambda \boldsymbol{\alpha} \tag{II.3.23}$$

$$\boldsymbol{\rho} = \left(\mathbb{1} - z\mathbf{\Pi K}\right)^{-1} z\boldsymbol{\alpha} \tag{II.3.24}$$

### II.3.3 Brief Notes on how we can take the Thermodynamic limit

For $\Phi$ a stable regular pair potential and $z$ a complex number satisfying $|z| < e^{-2\beta B - 1} C(\beta)^{-1}$. The grand canonical partition function has no zero in this range of $z$. If we define the correlation functions $\rho_\Lambda(x)_n$ as above for $z$ in this region, then we have infinite volume correlation functions $\rho(x)_n$ and a positive decreasing function $\epsilon$, such that:

$$\lim_{\lambda \to \infty} \epsilon(\lambda) = 0 \tag{II.3.25}$$

$$\|\rho_\Lambda(x)_n - \rho(x)_n\| \leq \xi^n \epsilon(\lambda) \tag{II.3.26}$$

where $\lambda$ is the minimum distance of $x_1, \cdots, x_n$ to the boundary of $\Lambda$.

**Theorem II.3.1** (The Thermodynamic Limit)**.** *If $\Lambda \to \infty$ in the sense of van Hove, then the following limits exist:*

$$\lim_{\Lambda \to \infty} \beta^{-1} V(\Lambda)^{-1} \log(\Xi_\Lambda(z, \beta)) =: P \tag{II.3.27}$$

$$\lim_{\Lambda \to \infty} z \frac{d}{dz} V(\Lambda)^{-1} \log \Xi_\Lambda(z, \beta) =: \rho \tag{II.3.28}$$

### II.3.4 The Relationship to the virial Expansion

We have the *virial expansion*:

$$\beta P = \sum_{n=1}^{\infty} c_n \rho^n \tag{II.3.29}$$

and wish to gain bounds on the coefficients $c_n$. In order to do so, we use a Cauchy integral representation of the $c_n$:

$$c_n = \frac{1}{2\pi i} \oint_C \frac{\mathrm{d}z}{nz\rho^{n-1}} \tag{II.3.30}$$

Providing that $C$ is a circle of radius less than $e^{-2\beta B-1}C(\beta)^{-1}$ about the origin in the complex $z$-plane. The above formula is fully derived in Section IV.2.

We use the operator norms from the Kirkwood Salsburg equations in order to achieve a bound on $|\rho - z|$:

$$\begin{aligned} |\rho - z| &\leq \xi \|\rho - z\boldsymbol{\alpha}\|_\xi \\ &\leq |z| \sum_{l=1}^{\infty} |z|^l \|\boldsymbol{\Pi K}\|_\xi^l \\ &= \frac{|z|^2 \|\boldsymbol{\Pi K}\|_\xi}{1 - |z| \|\boldsymbol{\Pi K}\|_\xi} \\ &\leq \frac{|z|^2}{e^{-2\beta B-1}C(\beta)^{-1} - |z|} \end{aligned} \tag{II.3.31}$$

and hence, via the reverse triangle inequality, we have a lower bound on $|\rho|$:

$$|\rho| \geq |z| - \frac{|z|^2}{e^{-2\beta B-1}C(\beta)^{-1} - |z|} \tag{II.3.32}$$

We need to maximise the right hand side of this inequality in terms of $|z|$, but still within the circle $C$. This maximum occurs at:

$$|z| = e^{-2\beta B-1}C(\beta)^{-1}\left(1 - \frac{1}{\sqrt{2}}\right) \tag{II.3.33}$$

giving us:

$$|\rho| \geq (3 - 2\sqrt{2})e^{-2\beta B-1}C(\beta)^{-1} \tag{II.3.34}$$

If we use this in the integral for $c_n$, then we achieve the bound:

$$|c_n| \leq \frac{1}{n}\left(e^{2\beta B+1}C(\beta)\frac{1}{(3 - 2\sqrt{2})}\right)^{n-1} \tag{II.3.35}$$

This gives the radius of convergence of the virial expansion ($\mathcal{R}_{\mathrm{Vir}}$) as:

$$\mathcal{R}_{\mathrm{Vir}} \geq e^{-2\beta B-1}C(\beta)^{-1}(3 - 2\sqrt{2}) \tag{II.3.36}$$

## II.4 Canonical Ensemble Calculations

The work by Pulvirenti and Tsagkarogiannis [PT12], gives a different point of view in understanding the virial expansion. They start from the Canonical Ensemble and then take the thermodynamic limit. The advantage of this method is that it avoids the need to invert power series in order to obtain a virial expansion from the cluster expansion. It also appears to be a much more natural approach, since we already have everything written in terms of the number of particles and volume, whose ratio is the density. There is no fugacity parameter in this approach and so no need to take an indirect route. They obtain Kotecký Preiss conditions on the convergence of the expansion and understand the relationship of what they obtain to what is obtained by the work of Mayer. Morais and Procacci [MoPr13] develop these ideas further and introduce bounds on the virial coefficients. They adopt the approach of an abstract polymer gas as in Section I.4 and use the tree-graph identity of Penrose [Pen67], to obtain improved coefficient bounds and the same bound on the radius of convergence as explained in subsection IV.5.1. They indicate that the canonical expansion method is just as effective as the inversion method and the bounds on coefficients indicate the possibility of improvements in this direction.

### II.4.1 The Model

In the Canonical Ensemble, we have a configuration of $N$ particles in a box $\Lambda$, interacting via a stable and tempered pair potential $\Phi : \mathbb{R}^D \to \mathbb{R}$. The boundary conditions are assumed to be periodic, that is we can understand the 'periodic' potential $\Phi^{\text{per}}(q_i, q_j)$, by splitting $\mathbb{R}^D$ into boxes of size $\Lambda$ and adding all of the interactions. In the case of $\Lambda$ being a box of side length $L$, we obtain:

$$\Phi^{\text{per}}(q_i, q_j) := \sum_{n \in \mathbb{Z}^D} \Phi(q_i - q_j + nL) \tag{II.4.1}$$

Stability is that $\exists B \geq 0$, such that for all $N$ and all $q_1, \cdots, q_N$:

$$\sum_{1 \leq i < j \leq N} \Phi(q_i - q_j) \geq -BN \tag{II.4.2}$$

Temperedness is that we assume the integral:

$$C(\beta) := \int_{\mathbb{R}^D} |e^{-\beta \Phi(q)} - 1| \mathrm{d}^D q \tag{II.4.3}$$

is convergent for some $\beta > 0$, and hence $\forall \beta > 0$.

The canonical partition function of this system is:

$$Z_N := \frac{1}{N!}\int_{\Lambda^N} \mathrm{d}(\mathbf{q})_N e^{-\beta H_\Lambda(\mathbf{q})} \tag{II.4.4}$$

where

$$H_\Lambda(\mathbf{q}) := \sum_{1\le i<j\le N} \Phi^{\mathrm{per}}(q_i, q_j) \tag{II.4.5}$$

Given a $\rho > 0$ (the *density*), we define the thermodynamic free energy as:

$$f_\beta(\rho) := \lim_{|\Lambda|,N\to\infty\ ,\ \frac{N}{|\Lambda|}\to\rho} f_{\beta,\Lambda}(N) \tag{II.4.6}$$

where

$$f_{\beta,\Lambda}(N) := -\frac{1}{\beta|\Lambda|}\ln Z_N \tag{II.4.7}$$

The main result of the paper is:

**Theorem II.4.1** (Convergence and Interpretation of the Free Energy Expansion - [PT12])**.** *For $\rho C(\beta)e^{2\beta B+\alpha+1} < 1$, where $\alpha$ is a constant related to the Kotecký Preiss condition used, we have*

$$\frac{1}{|\Lambda|}\ln Z_N = \ln\frac{|\Lambda|^N}{N!} + \frac{N}{|\Lambda|}\sum_{n\ge 1} F_{N,\Lambda}(n) \tag{II.4.8}$$

*where the coefficients $F_{N,\Lambda}(n)$, $n \ge 1$, satisfy:*

$$|F_{N,\Lambda}(n)| \le Ce^{-cn} \tag{II.4.9}$$

*For constants $C$ and $c > 0$, which are independent of $N$ and $\Lambda$.*

*We can interpret these coefficients as follows:*

$$F_{N,\Lambda}(n) = \frac{1}{n+1}P_{N,|\Lambda|}(n)B_{\beta,\Lambda}(n) \tag{II.4.10}$$

*defined by:*

$$P_{N,|\Lambda|}(n) := \frac{(N-1)\cdots(N-n)}{|\Lambda|^n} \tag{II.4.11}$$

$$B_{\beta,\Lambda}(n) := \frac{|\Lambda|^n}{n!}\sum_{\{V_1,\cdots,V_n\}\in[N]^n} \phi^T(V_1,\cdots,V_n)\zeta_\Lambda(V_1)\cdots\zeta_\Lambda(V_n) \tag{II.4.12}$$

*where the $\phi$ and $\zeta$ are explained in the context of a polymer model below.*

*These have thermodynamic limits:*

$$\lim_{N,|\Lambda|\to\infty\ ,\ \frac{N}{|\Lambda|}=\rho} P_{N,|\Lambda|}(n) = \rho^n \quad \textit{and} \quad \lim_{\Lambda\to\infty} B_{\beta,\Lambda}(n) = \beta_n \tag{II.4.13}$$

*for all $n \geq 1$, where $\beta_n$ are the irreducible coefficients of Mayer:*

$$\beta_n := \frac{1}{n!} \sum_{\substack{g\in\mathcal{B}[n+1] \\ supp\ g\ni\{1\}}} \int_{(\mathbb{R}^D)^n} \prod_{\{i,j\}\in E(g)} (e^{-\beta\Phi(q_i-q_j)} - 1)\mathrm{d}(\mathbf{q})'_{n+1} \quad q_1 = 0 \tag{II.4.14}$$

The approach is first to divide the canonical partition function into the ideal and the interacting parts:

$$Z_N = Z_N^{\text{ideal}} Z_N^{\text{int}} \tag{II.4.15}$$

where:

$$Z_N^{\text{ideal}} := \frac{|\Lambda|^N}{N!} \quad \text{and} \quad Z_N^{\text{int}} := \int_{\Lambda^N} \frac{\mathrm{d}q_1}{|\Lambda|} \cdots \frac{\mathrm{d}q_N}{|\Lambda|} e^{-\beta H_\Lambda(\mathbf{q})} \tag{II.4.16}$$

Of course $Z_N^{\text{ideal}}$ gives us the first term easily and we have to concentrate on the interaction part. The first stage is to use the idea by Mayer to write the interaction part of the partition function as a sum over graphs with edge-weights the Mayer $f$-functions [MMay40]:

$$e^{-\beta H_\Lambda(\mathbf{q})} = \prod_{1\leq i<j\leq N} (1 + f_{i,j}) = \sum_{g\in\mathcal{G}[N]} \prod_{\{i,j\}\in E(g)} f_{i,j} \tag{II.4.17}$$

The next stage is to use this formulation to draw a parallel with abstract polymer models. What is required to draw such a connection is an *incompatibility relation*, which is conveyed in the paper of Procacci [Pro07].

The support of a graph, denoted supp $g$, is the set of its vertices. We call a graph connected if there exists a path from each vertex in the graph to any other. We define the incompatibility relation as follows: $g$ and $g'$ are incompatible if supp $g \cap$ supp $g' \neq \emptyset$. They are compatible if they have empty intersection. This obeys symmetry trivially and is reflexive unless our graph is the empty graph, but we will not allow for these in the formulation (the empty graph is not connected). We call the cardinality of supp $g$, $|g|$. Any graph $g \in \mathcal{G}[N]$ is equivalent to the pairwise compatible (non-ordered) collection of its connected components: i.e. $g = \{g_1, \cdots, g_k\}$ for some $k$. So we write the canonical partition function as that of a

weighted polymer:

$$Z_N^{\text{int}} := \sum_{\{g_1,\cdots g_k\}_\sim} \prod_{i=1}^{k} \tilde{\zeta}_\Lambda(g_i) = \sum_{\{V_1,\cdots,V_k\}_\sim} \prod_{i=1}^{k} \zeta_\Lambda(V_i) \tag{II.4.18}$$

where the weights are defined as:

$$\zeta_\Lambda(V) := \sum_{g\in\mathcal{C}[V]} \tilde{\zeta}_\Lambda(g), \qquad \tilde{\zeta}_\Lambda(g) := \int_{\Lambda^{|g|}} \prod_{i\in \text{supp } g} \frac{\mathrm{d}q_i}{|\Lambda|} \prod_{\{i,j\}\in E(g)} f_{i,j} \tag{II.4.19}$$

where the empty product is defined to be 1. The expression for the logarithm of this function is:

$$\ln Z_N = \sum_{n=0}^{\infty} \frac{1}{n!} \sum_{(V_1,\cdots,V_n)\in[N]^n} \phi^T(V_1,\cdots,V_n)\zeta_\Lambda(V_1)\cdots\zeta_\Lambda(V_n) \tag{II.4.20}$$

where:

$$\phi^T(V_1,\cdots,V_n) = \begin{cases} 1 & \text{if } n=1 \\ \sum\limits_{\substack{g\in\mathcal{C}[n] \\ g\subset G(V_1,\cdots,V_n)}} (-1)^{|E_g|} & \text{if } n\geq 2 \end{cases} \tag{II.4.21}$$

where $G(V_1,\cdots V_n)$ is the graph on vertex set $[n]$ with edge set $E_{G(V_1,\cdots,V_n)} = \{\{i,j\}\subset[n] | V_i\cap V_j\neq\emptyset\}$ We know that $\ln Z_N$ can be written as an absolutely convergent power series for all complex activities $\zeta_\Lambda(V)$, as soon as:

$$\sup_{i\in[N]} \sum_{\substack{V\subset[N]|i\in V \\ |V|\geq 2}} |\zeta_\Lambda(V)|e^{a|V|} \leq e^a - 1 \tag{II.4.22}$$

This is analogous to the Kotecký Preiss condition in Section I.4. The activity $\zeta_\Lambda(V)$ depends only on $|V|$ and so we rewrite the condition in the convenient form:

$$\sum_{m=2}^{N} e^{am} C_m^\rho \leq e^a - 1 \tag{II.4.23}$$

where

$$C_m^\rho = |\zeta_m| \binom{N-1}{m-1} \tag{II.4.24}$$

We can identify the relationship between $\zeta_m$ and the cluster coefficients $b_m$, by:

$$\zeta_m = \frac{b_m(\beta\Lambda)m!}{V^{m-1}} \tag{II.4.25}$$

and so the cluster expansion bounds above give us the suitable bound on the radius of convergence.

The inspiration of the Kotecký Preiss like condition is from the paper by Bovier and Zahradník [BoZa00], which gives a simple inductive proof of Dobrushin's criterion.

Define:

$$L = L(\delta) := \sup_{x \in (0,\delta)} \left( \frac{-\ln(1-x)}{x} \right) = \frac{-\ln(1-\delta)}{\delta} \tag{II.4.26}$$

for $\delta$ small.

**Theorem II.4.2** (Bovier-Zahradník Condition)**.** *Suppose there exist two non-negative functions $a, c : \Gamma \to \mathbb{R}$ and $\forall \gamma \in \Gamma$*

$$|w(\gamma)| e^{a(\gamma)} \leq \delta \tag{II.4.27}$$

*For some small $\delta > 0$.*

*Assume also for any polymer $\gamma'$:*

$$\sum_{\gamma \nsim \gamma'} |w(\gamma)| e^{a(\gamma)+c(\gamma)} \leq \frac{1}{L} a(\gamma') \tag{II.4.28}$$

*Then we have the following bound:*

$$\sum_{I | I(\gamma') \geq 1} |c_I W^I| e^{\sum_{\gamma \in supp I} I(\gamma) c(\gamma)} \leq L |w(\gamma')| e^{a(\gamma')+c(\gamma')} \tag{II.4.29}$$

The Bovier-Zahradník condition is applied to the relevant polymer model.

We have that

$$F_{N,\Lambda}(n) := \frac{1}{n+1} \sum_{\substack{I | \text{ supp } I \ni 1 \\ |I| = n+1}} C_I \zeta_\Lambda^I \tag{II.4.30}$$

where $|I| := |\cup_{V \in \text{ supp } I} V| \leq \|I\|$. The Bovier Zahradník bound then gives us that $F_{N,\Lambda}(n)$ is uniformly bounded for all $N, \Lambda$ as well as absolutely summable over $n$. The bound gives us:

$$|F_{N,\Lambda}(n)| \leq \frac{e^{-cn}}{n+1} \sum_{\substack{I | \text{ supp } I \ni 1 \\ |I| = n+1}} \left| C_I \zeta_\Lambda^I \right| e^{cn} \leq e^{-cn} L e^{\alpha} \tag{II.4.31}$$

The paper obtains the two-connected graph expression for the coefficients $F_{\Lambda,N}(n)$

in the thermodynamic limit. This relies on the block-factorisation property of the weights $\zeta_\Lambda(V)$, when understood as depending on a graph $g \in \mathcal{C}[n]$. The key point is that the product structure mentioned in the paper gives that no mixed terms appear. For connected graphs the block factorisation property gives that only terms with two-connected graphs contribute.

## II.5 Conclusions & Open Questions

This chapter has introduced Mayer's theory of virial expansions and given the classical approach of Kirkwood Salsburg equations and the inversion of cluster expansions. In Chapter IV, improvements of this approach are given, yet it is still unclear whether this is an optimal approach, since, especially in the positive potential case, one is strongly dependent on the unphysical singularity of the cluster expansion on the negative real axis.

The other key approach indicated is that of working in the Canonical Ensemble instead of using the inversion. The paper of Pulvirenti and Tsagkarogiannis [PT12] indicates that we may derive all of the classical results of two-connected graphs and similar bounds through this approach. The further paper by Morais and Procacci [MoPr13] captures the bounds of Lebowitz and Penrose [LePen64]. This is also emphasised in Chapter IV. The paper indicates the possibility of improving such bounds and it is still an open question to make improved bounds via the Canonical Ensemble.

# Chapter III

# A Graphical Involution Solving the Puzzle of Mayer's Virial Expansion

The main content of this chapter is to give a direct proof for four identities arising from Mayer's theory of cluster and virial expansions for the one particle hard core gas and the continuum hard core gas or Tonks gas. This is to identify the combinatorial cancellations. The main relationships are:

$$\sum_{g\in\mathcal{C}[n]}(-1)^{e(g)}=(-1)^{n-1}(n-1)! \tag{III.0.1}$$

$$\sum_{g\in\mathcal{B}[n]}(-1)^{e(g)}=-(n-2)! \tag{III.0.2}$$

for the one particle hard core gas, where $e(g)$ denotes the number of edges in the graph $g$.

Let the polytope corresponding to the graph $g$ be defined as:

$$\Pi_g:=\{(\mathbf{x})_{[2,N]}\in\mathbb{R}^{N-1}|\,|\mathbf{x}_i-\mathbf{x}_j|<1\;\forall\{i,j\}\in E(g)\} \tag{III.0.3}$$

with $x_1=0$. We have:

$$\sum_{g\in\mathcal{C}[n]}(-1)^{e(g)}\,\mathrm{Vol}(\Pi_g)=(-1)^{n-1}n^{n-1} \tag{III.0.4}$$

$$\sum_{g\in\mathcal{B}[n]}(-1)^{e(g)}\,\mathrm{Vol}(\Pi_g)=-n(n-2)! \tag{III.0.5}$$

for the Tonks gas. The explanation in this form for the connected identities is given by Bernardi [Ber08] and the two-connected case is original work. The explanation by Bernardi involves indicating that (III.0.1) can be resolved in terms of enumerating increasing trees and (III.0.4) can be resolved in terms of enumerating rooted trees.

The work of Mayer [MMay40] introduced an important connection between the subjects in the form of representing coefficients of the important expansions of statistical mechanics - the partition function, the cluster expansion and the virial expansion - in terms of weighted graphs of a particular type - simple graphs, connected graphs and two-connected graphs respectively. Most recently, the combinatorial species of structure framework introduced by Joyal [Joy81], has been applied to statistical mechanics through the work of Ducharme, Labelle and Leroux [DLL07, Ler04], Kaouche and Leroux [KaLe08] and Faris [Far10], in order to give an interpretation to these important connections. Useful developments of the subject of combinatorial species of structure can be found in the book by Bergeron Labelle and Leroux [BLL98] and Flajolet and Sedgewick [FlSe09]. Such connections between the two subjects provide mutual exchanges. The main exchange that this chapter focuses upon are the combinatorial identities indicated in the paper [DLL07] afforded to us by using the combinatorial connection with two simplistic statistical mechanical models, namely the hardcore one-particle gas and the hardcore continuum gas.

The chapter is organised so that, in Section III.1, the two models of the one particle hard core gas and the Tonks gas are explained alongside what Mayer's theorems give in these cases. The bc-tree used in the proof is also introduced in Section III.2 and the explanation of polytopes and their decomposition into simplicies, attributed to Lass is indicated in Section III.3, which is instrumental to the proof in the Tonks gas case. A short discussion of the proof by Bernardi [Ber08] is given in Section III.4. Section III.5 presents the the combinatorial structures which describe the particular coefficients and gives a discussion on the meaning of the results. Sections III.6 and III.7 give the proofs of the one particle hard core and the Tonks gas case respectively. Section III.9 ends with what is hoped to be future directions and context of these combinatorial interpretations and ideas of future work.

## III.1 The Two Models from Statistical Mechanics

Recall the graph weights from Section I.2:

$$w(g) = \prod_{\{i,j\} \in E(g)} f_{i,j} \tag{III.1.1}$$

$$W(g) := \prod_{i \in V(g)} \left( \int_V dq_i \right) w(g) \tag{III.1.2}$$

which are equations (I.2.8) and (I.2.10), respectively. The pressure function can be written in terms of connected graphs:

$$\beta P = \mathcal{C}_W(z) = \sum_{n=1}^{\infty} \frac{z^n}{n!} \sum_{g \in \mathcal{C}[n]} W(g) \tag{III.1.3}$$

This is the content of Mayer's First Theorem [MMay40]. The density $\rho$ is:

$$\rho = z \frac{\partial}{\partial z} \beta P = \mathcal{C}^{\bullet}_W(z) \tag{III.1.4}$$

where $\mathcal{C}^{\bullet}$ denotes a rooted connected graph. From Mayer's Second Theorem [MMay40] found in Section II.2 or by the Dissymmetry Theorem [DLL07], we are able to obtain a series expansion for pressure in terms of density, in which the coefficients are, up to a prefactor, the $W$-weighted two-connected graphs.

$$\beta P = \rho - \sum_{n=2}^{\infty} \frac{(n-1)\rho^n}{n!} \sum_{g \in \mathcal{B}[n]} W(g) \tag{III.1.5}$$

### III.1.1 One Particle Hard Core Gas

The potential for a one-particle hard core gas is:

$$\Phi_{\text{1PHG}}(q_i, q_j) = \infty \tag{III.1.6}$$

so that $\exp(-\beta \Phi_{\text{1PHG}}(q_i, q_j)) = 0$ and $f(q_i, q_j) = -1$. The grand canonical partition function is

$$\Xi(z) = 1 + z \tag{III.1.7}$$

The statistical mechanical relationships give pressure and density as:

$$\beta P = \log(1 + z) \tag{III.1.8}$$

$$\rho = \frac{z}{1 + z} \tag{III.1.9}$$

We may invert (III.1.9), to obtain:

$$z = \frac{\rho}{1 - \rho} \tag{III.1.10}$$

and substitute for $z$ in (III.1.8), to obtain:

$$\beta P = -\log(1-\rho) \tag{III.1.11}$$

The two series expansions derived from statistical mechanics are:

$$\beta P = \sum_{n=1}^{\infty} \frac{(-1)^{n-1} z^n}{n} \tag{III.1.12}$$

$$\beta P = \sum_{n=1}^{\infty} \frac{\rho^n}{n} \tag{III.1.13}$$

If we compare these two power series with (III.1.3) and (III.1.5) respectively, using the graph weight $W(g) = (-1)^{e(g)}$, we obtain:

$$\sum_{g \in \mathcal{C}[n]} (-1)^{e(g)} = (-1)^{n-1}(n-1)! \tag{III.1.14}$$

$$\sum_{g \in \mathcal{B}[n]} (-1)^{e(g)} = -(n-2)! \tag{III.1.15}$$

We emphasise here that we hit the lower bound given by Groeneveld in Section I.7.1. We have that $|b_n| = \frac{1}{n}$ and noting that $|2b_2| = 1$, we have the lower bound for the cluster coefficients for positive potentials given by Groeneveld.

### III.1.2 Continuum Hard Core Gas - Tonks Gas

For a continuum hard core gas in one dimension with diameter 1, the potential is:

$$\Phi_{\mathrm{TG}}(q_i, q_j) = \begin{cases} \infty & \text{if} \quad |q_i - q_j| < 1 \\ 0 & \text{otherwise} \end{cases} \tag{III.1.16}$$

The exponential and Mayer $f$-functions are:

$$\exp(-\beta \Phi_{\mathrm{TG}}(q_i, q_j)) = \begin{cases} 0 & \text{if} \quad |q_i - q_j| < 1 \\ 1 & \text{otherwise} \end{cases} \tag{III.1.17}$$

$$f(q_i, q_j) = \begin{cases} -1 & \text{if} \quad |q_i - q_j| < 1 \\ 0 & \text{otherwise} \end{cases} \tag{III.1.18}$$

We therefore have the graph weight:

$$w(g) = (-1)^{e(g)} \int_{\mathbb{R}^{n-1}} \prod_{(i,j)\in E_g} \chi(|x_i - x_j| < 1)\, \mathrm{d}x_2 \cdots \mathrm{d}x_n \tag{III.1.19}$$

where $x_1 = 0$ and $\chi$ is the indicator function.

In [DLL07], this is interpreted as the volume of a convex polytope $\Pi_g$ in $\mathbb{R}^{n-1}$, defined in (III.0.3).

Hence the graph weight may be written as:

$$W(g) = (-1)^{e(g)} \operatorname{Vol}(\Pi_g) \tag{III.1.20}$$

We may derive the cluster and virial expansions for the Tonks gas in the general case where we have a hard core radius of $r_{h,c}$ and then specialise to the case where $r_{h,c} = 1$. The easiest way to make computations is via the Canonical Ensemble. For ease of notation, we write $\exp(-\beta\Phi_{\mathrm{TG}}(q_i, q_j)) = e_{i,j}$. We enclose our rods in the space $[0, L]$ to obtain finite integrals and the canonical partition function is:

$$Z_N = \int_{\frac{r_{hc}}{2}}^{L-\frac{r_{hc}}{2}} \mathrm{d}q_1 \cdots \int_{\frac{r_{hc}}{2}}^{L-\frac{r_{hc}}{2}} \mathrm{d}q_N \prod_{i=1}^{N-1} e_{i,i+1} \tag{III.1.21}$$

$$= \int_{\frac{r_{hc}}{2}}^{L+\frac{r_{hc}}{2}-Nr_{hc}} \mathrm{d}q_1 \int_{q_1+r_{hc}}^{L+\frac{r_{hc}}{2}-Nr_{hc}+r_{hc}} \mathrm{d}q_2 \cdots \int_{q_{N-1}+r_{hc}}^{L+\frac{r_{hc}}{2}-Nr_{hc}+(N-1)r_{hc}} \mathrm{d}q_N \tag{III.1.22}$$

We make the change of variables: $\omega_k = q_k - \left(k + \frac{1}{2}\right) r_{hc}$:, to get

$$Z_N = \int_0^{L-Nr_{hc}} \mathrm{d}\omega_1 \int_{\omega_1}^{L-Nr_{hc}} \mathrm{d}\omega_2 \cdots \int_{\omega_{N-1}}^{L-Nr_{hc}} \mathrm{d}\omega_N \tag{III.1.23}$$

$$= \int_0^{L-Nr_{hc}} \mathrm{d}\omega_1 \cdots \int_{\omega_{k-1}}^{L-Nr_{hc}} \mathrm{d}\omega_k \frac{(L - Nr_{hc} - \omega_k)^{N-k}}{N-k)!} = \int_0^{L-Nr_{hc}} \mathrm{d}\omega_1 \frac{(L - Nr_{hc} - \omega_1)^{N-1}}{(N-1)!} \tag{III.1.24}$$

So:

$$Z_N = \frac{(L - Nr_{hc})^N}{N!} \tag{III.1.25}$$

In the Canonical Ensemble, then we may derive the equation of state, by using the

relation:

$$\beta P = \frac{\partial}{\partial L}(\ln Z_N) = \frac{N}{L - N r_{hc}} = \frac{\rho}{1 - \rho r_{hc}} \tag{III.1.26}$$

taking $\frac{N}{L} = \rho$.

We use the relationship $\rho = z\frac{d}{dz}\beta P$ and (III.1.26), to obtain a differential equation for $\beta P = \chi$ in terms of $z$:

$$\begin{aligned}
\chi &= \frac{z\frac{d\chi}{dz}}{1 - z\frac{d\chi}{dz}r_{hc}} \\
\chi &= z\frac{d\chi}{dz}(1 + r_{hc}\chi) \\
\int \frac{\mathrm{d}z}{z} &= \int \frac{1}{\chi} + r_{hc}\mathrm{d}\chi \\
\ln z &= \ln \chi + r_{hc}\chi \\
r_{hc}z &= r_{hc}\chi e^{r_{hc}\chi}
\end{aligned} \tag{III.1.27}$$

We make contact here with the Lambert $W$-function, which is defined as the inverse of the mapping $z \mapsto ze^z$. This inverse has many branch cuts, but we take the 0-branch cut, which is real. We may then use the Lambert $W$-function in (III.1.27), to obtain the expression for pressure:

$$\beta P = \frac{1}{r_{\mathrm{h\,c}}} W(r_{\mathrm{h\,c}} z) \tag{III.1.28}$$

More details are given about the Lambert $W$-function in Section IV.4. The key point here is that the radius of convergence for $W(r_{\mathrm{h\,c}}z)$ as a power series in $z$, is $\frac{1}{r_{\mathrm{h\,c}}e}$ and $|2b_2|$ is simply $r_{\mathrm{h\,c}}$. We therefore obtain the lower bound on the radius of convergence for positive potentials detailed by Groeneveld, given in Section I.7.1.

If we now specialise to $r_{hc} = 1$, we have the expressions:

$$\beta P = W(z) = \sum_{n=1}^{\infty} \frac{(-1)^n n^{n-1} z^n}{n!} \tag{III.1.29}$$

$$\beta P = \frac{\rho}{1 - \rho} = \sum_{n=1}^{\infty} \rho^n \tag{III.1.30}$$

If we compare these to the results of Mayer's First and Second Theorems

(III.1.3) and (III.1.5), we obtain the combinatorial relationships:

$$\sum_{g\in\mathcal{C}[n]} (-1)^{e(g)} \operatorname{Vol}(\Pi_g) = (-1)^{n-1} n^{n-1} \tag{III.1.31}$$

$$\sum_{g\in\mathcal{B}[n]} (-1)^{e(g)} \operatorname{Vol}(\Pi_g) = -n(n-2)! \tag{III.1.32}$$

## III.2 The Block Cutpoint Tree

This section briefly introduces the notion of the block cutpoint tree used in the proof.

- An *articulation point* in a connected graph is a vertex, which when it and its incident edges are removed, renders the graph disconnected. A synonym that is frequently used is a *cutpoint*.
- A *two-connected graph* is a connected graph without articulation points.
- A *block* is a maximal two-connected subgraph of a connected graph. Maximal in terms of edges and vertices it includes.

The block cutpoint tree (bc-tree) associated to a connected graph $g$ is a (bipartite) graph where the vertices represent the articulation points and the blocks in a connected graph. An edge, between an articulation point and a block, is present in this graph, when an articulation point is contained in a block. It is a tree, since if there were a cycle in this graph then the cycle itself would have been a block. An example of a block cutpoint tree is shown in Figure III.1.

**Definition** (The Centre of a Tree)**.** *To define the centre of a tree formally, we define first the eccentricity $\varepsilon(v)$ of a vertex $v$ as the minimal graph distance of $v$ to a leaf. The centre of a tree is the collection of vertices at which the maximum eccentricity is attained. This can either be two neighbouring vertices or a single vertex. In the former case, we often call the edge between the vertices the centre of the tree.*

**Remark 18** (An Algorithmic Interpretation of the Centre of the Tree)**.** *One can apply the function $f : \mathfrak{a} \to \mathfrak{a}$, which for any given tree, removes all leaves and the edges corresponding to the leaves. Formally, we can write this as:*

$$f : (V(\tau), E(\tau)) \mapsto (V(\tau) \setminus L, E(\tau) \setminus (L \times V(\tau))) \tag{III.2.1}$$

*where $L := \{i \in V(\tau) | \deg(i) = 1\}$, the collection of leaves.*

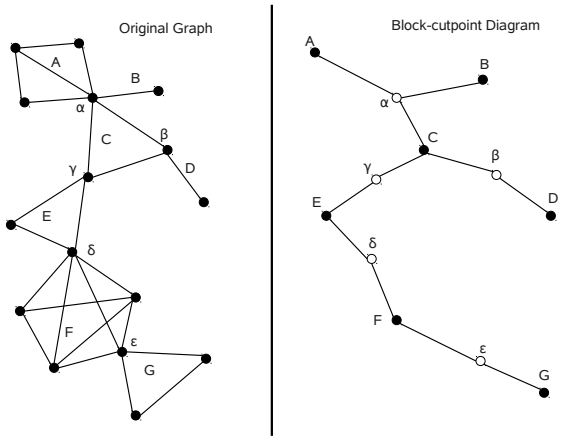



Figure III.1: An example of a bc-tree

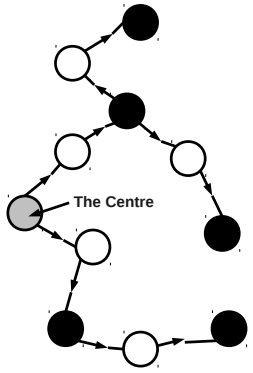



Figure III.2: The associated digraph

*Repeated application of $f$, gives a sequence of trees, $(f^n(\tau))_{n\in\mathbb{N}_0}$ which becomes constant either when we have a single vertex or the empty graph. In the case of the single vertex, this is the centre of the tree. For the empty graph, the penultimate step will have been two vertices and an edge and this edge or the pair of vertices is defined as the centre.*

A bc-tree is bipartite with all leaves in one set (the blocks). It therefore has a unique centre, since the eccentricity of the articulation points will be odd and the eccentricity of the blocks will be even so two neighbours cannot have the same maximum eccentricity.

**Lemma III.2.1** (Simple bc-tree Identity)**.** *For any connected graph, we have the identity:*

$$\sum_{i\in I}(k_i - 1) = n - 1 \tag{III.2.2}$$

*where $I$ is the label set for the blocks $k_i$ indicates the number of vertices in a block and $n$ is the total number of vertices in the connected graph.*

*Proof.* The key idea is to indicate what vertex we omit inside each block on the left hand side of (III.2.2). The digraph explained below gives an (essentially) unique prescription of the missing vertex in each block and in which block an articulation point is counted.

In the bc-tree, we can give all edges the direction away from the central vertex, as can be seen in Figure III.2. In this digraph, we have two types of directed edge $(B, a)$ and $(a, B)$, where $a$ indicates an articulation point and $B$ a block. The

arrow points from the first entry to the second entry. Since there is a unique path from the centre to every other vertex, every vertex has precisely one edge in which they are the second entry.

There are two key cases:

i) **The centre is an articulation point**

For a block, $B$, the unique vertex we neglect on the left hand side of (III.2.2) is the articulation point, $a$, where $(a, B)$ is the directed edge in the digraph.

Every articulation point, $\alpha$, except the centre appears in an edge $(\beta, \alpha)$, for which it is the second entry, meaning it is enumerated in the left hand side of (III.2.2) in precisely one block. The central articulation point is the only neglected vertex, which gives the right hand side of (III.2.2).

ii) **The centre is a block**

In this case every block, except the centre, can be given the prescription as for the first case. For the central block, we can choose precisely one of its neighbours to neglect. All articulation points in this case have an edge in which they are the second entry and so are counted, excepting the articulation point identified by the central block. Therefore, we have (III.2.2).

□

## III.3 Polytopes and Simplices

In [DLL07] there is a decomposition of the polytope $\Pi_g$ into simplices, which is used in the proof given in [Ber08]. Bernardi in [Ber08] gives an explanation of this procedure. This section explains this procedure.

Consider $(\mathbf{x})_{[2,n]} \in \mathbb{R}^{n-1}$ and let $h_i$ be the integer part of $x_i$ and $0 \leq w_i < 1$ be the fractional part such that $h_i + w_i = x_i$. Let $\sigma : [2, n] \to [2, n]$ be a bijection. We may define the simplex $\pi(\mathbf{h}, \sigma)$, by the set of $\mathbf{x}$ with integer part $\mathbf{h}$ and whose fractional parts satisfy: $w_{\sigma(2)} < w_{\sigma(3)} < \cdots < w_{\sigma(n)}$. This simplex has volume $\frac{1}{(n-1)!}$.

The condition $|x_i - x_j| < 1$ is equivalent to $h_i - h_j \in \{0, \text{sign}(w_j - w_i)\}$. We therefore have that $\pi(\mathbf{h}, \sigma) \subset \Pi_g$ if and only if for all $(i, j) \in g$, we have that $h_i - h_j \in \{0, \text{sign}(\sigma^{-1}(j) - \sigma^{-1}(i))\}$ with $h_1 = 0$ and $\sigma(1) = 1$.

**Lemma III.3.1.** *For any graph $g \in \mathcal{G}[n]$, the value $(n-1)!\,\text{Vol}(\Pi_g)$ counts the pairs $\mathbf{h} \in \mathbb{Z}^{n-1}$ and $\sigma \in S_{n-1}$ such that $\pi(\mathbf{h}, \sigma)$ is a subpolytope of $\Pi_g$.*

We may rearrange the sums over connected or two-connected graphs of the graph weights by first casting the sum as a sum over the pairs $(\mathbf{h}, \sigma)$ and symmetrising the weight over isomorphic graphs. The symmetrisation procedure can be understood by considering a permutation $\sigma$ of $[2, n]$ and defining for any vector $\mathbf{h} = (h_2, \cdots, h_n) \in \mathbb{Z}^{n-1}$, $\sigma(\mathbf{h}) = (h_{\sigma(2)}, \cdots, h_{\sigma(n)})$. For any graph $g$ with labels in $[n]$, the graph $\sigma(g)$ is the graph, with the same vertex set and satisfies $(\sigma(i), \sigma(j)) \in E(\sigma(g)) \iff (i, j) \in E(g)$.

Observe that $\pi(\mathbf{h}, \sigma) \subset \Pi_g$ if and only if $\pi(\sigma^{-1}(\mathbf{h}), \mathrm{Id}) \subset \Pi_{\sigma(g)}$ for any permutation $\sigma$ of $[2, n]$. We can see this equivalence, by rewriting $\mathbf{w} = \sigma^{-1}(\mathbf{h})$ and $h = \sigma(g)$ to see that the latter statement can be cast as $\pi(\mathbf{w}, \mathrm{Id}) \subseteq \Pi_h$. This means in terms of the entries in vector $\mathbf{w}$ that $\forall (k, l) \in E(h)\ w_k - w_l \in \{0, \mathrm{sign}(l - k)\}$. Since $(i, j) \in E(g) \iff (\sigma(i), \sigma(j)) \in E(h)$, we may rewrite this as: $\forall (i, j) \in E(g)$ $w_{\sigma(i)} - w_{\sigma(j)} \in \{0, \mathrm{sign}(\sigma(j) - \sigma(i))\}$. We make the identification that $h_i = w_{\sigma(i)}$ to see that we get precisely the statement that $\pi(\mathbf{h}, \sigma) \subseteq \Pi_g$.

We let $\mathcal{H}$ denote either $\mathcal{C}$ or $\mathcal{B}$ and then we rewrite:

$$\begin{aligned}\sum_{\substack{\mathbf{h}\in\mathbb{Z}^{n-1}\, g\in\mathcal{H}[n]\\ \pi(\mathbf{h},\sigma)\subset\Pi_g}} (-1)^{e(g)} &= \sum_{\substack{\mathbf{h}\in\mathbb{Z}^{n-1}\, g\in\mathcal{H}[n]\\ \pi(\sigma^{-1}(\mathbf{h}),\mathrm{Id})\subset\Pi_{\sigma(g)}}} (-1)^{e(g)} \\ &= \sum_{\substack{\mathbf{h}\in\mathbb{Z}^{n-1}\, g\in\mathcal{H}[n]\\ \pi(\mathbf{h},\mathrm{Id})\subset\Pi_g}} (-1)^{e(\sigma^{-1}(g))} \\ &= \sum_{\substack{\mathbf{h}\in\mathbb{Z}^{n-1}\, g\in\mathcal{H}[n]\\ \pi(\mathbf{h},\mathrm{Id})\subset\Pi_g}} (-1)^{e(g)} \qquad \text{(III.3.1)}\end{aligned}$$

We may therefore, understand the weight as:

$$\begin{aligned}\sum_{g\in\mathcal{H}[n]} w(g) = \sum_{g\in\mathcal{H}[n]} (-1)^{e(g)} \mathrm{Vol}(\Pi_g) &= \frac{1}{(n-1)!} \sum_{\substack{\mathbf{h}\in\mathbb{Z}^{n-1}\sigma\in S_{n-1}\\ \text{such that } \pi(\mathbf{h},\sigma)\subset\Pi_g}} (-1)^{e(g)} \\ &= \sum_{\substack{\mathbf{h}\in\mathbb{Z}^{n-1}\, g\in\mathcal{H}[n]\\ \pi(\mathbf{h},\mathrm{Id})\subset\Pi_g}} (-1)^{e(g)} \qquad \text{(III.3.2)}\end{aligned}$$

We define the *centroid* of the vector $\mathbf{h}$, by $\bar{\mathbf{h}} = (\bar{h_1}, \cdots, \bar{h_n})$, where $\bar{h_i} = h_i + \frac{i-1}{n}$. We define $G_{\mathbf{h}}$ as the graph on $[n]$ where the edges are all pairs $(i, j)$ such that $|\bar{h_i} - \bar{h_j}| < 1$. We define $\mathcal{H}_{\mathbf{h}}[n] := \{g \in \mathcal{H}[n] | E(g) \cap E(G_{\mathbf{h}}) = E(g)\}$ where $\mathcal{H}$ can be replaced by $\mathcal{C}$ or $\mathcal{B}$.

The final sum indicates that we need to count pairs $\mathbf{h}$ and $g$ such that $\pi(\mathbf{h}, \mathrm{Id}) \subset \Pi_g$. That is that the centroid $\bar{\mathbf{h}} \in \Pi_g$, since $\bar{\mathbf{h}}$ is in the interior of

$\pi(\mathbf{h}, \mathrm{Id})$. This can be recast as: for $\bar{\mathbf{h}} \in \Pi_g$, we require that:

$$\forall (i,j) \in E(g) \; |\bar{h}_i - \bar{h}_j| < 1 \tag{III.3.3}$$

We can, therefore, rewrite our sum as:

$$\sum_{\mathbf{h} \in \mathbb{Z}^{n-1}} \sum_{g \in \mathcal{H}_{\mathbf{h}}} (-1)^{e(g)} \tag{III.3.4}$$

we can thus consider the total sum as first a sum over the subset of graphs $\mathcal{H}_{\mathbf{h}}$ for each $\mathbf{h}$ and add the results. This leads to considering separate $\Psi_{\mathbf{h}} : \mathcal{B}_{\mathbf{h}} \to \mathcal{B}_{\mathbf{h}}$ which are involutions and finding their fixed points.

## III.4 Bernardi's Interpretation

The main idea to see how the cancellations work is to construct an involution on the set of connected graphs, which changes the number of edges by at most one. This gives a pairing of most connected graphs to leave the unpaired graphs as being the only contribution to the sum. This has a connection to matroids, graph tree inequalities and the Penrose identity, which is explained in Chapter VII. The key idea is that this construction has generalisations, which make for useful applications to the cluster expansion. This is to motivate understanding the same perspective for two-connected graphs, as providing insight into virial coefficients.

The lexicographic order on edges in a graph is defined by:

$$(i,j) < (k,l) \text{ if } \begin{cases} & \min\{i,j\} < \min\{k,l\} \\ \text{or} & \min\{i,j\} = \min\{k,l\} \text{ and } \max\{i,j\} < \max\{k,l\} \end{cases} \tag{III.4.1}$$

Given a graph $g = (V, E)$, we define the graph $g^{>e} = (V, E^{>e})$, where $e \in V^{(2)}$ and $E^{>e} = E \cap \{y \in V^{(2)} | y > e\}$. We call an edge $(i,j)$ $g$-active, if there exists a path from $i$ to $j$ in $g^{>(i,j)}$. We let $e_g^\star$ be the greatest $g$-active edge, if there is one. This notion of being an active edge is connected to the active edges of matroids explained in the paper by Sokal [Sok05]. This is interpreted in Chapter VII.

Define the mapping $\Psi : \mathcal{C} \to \mathcal{C}$ by:

$$\Psi(g) = \begin{cases} g & \text{if there is no } g\text{-active edge} \\ g \oplus e_g^\star & \text{if we have a } g\text{-active edge} \end{cases} \tag{III.4.2}$$

The operation $\oplus$ means to add the edge if it is not present and remove it if it is.

What is left over after those which add or lose an edge cancel? We understand if $\Psi(g) \neq g$, then $(-1)^{e(g)} + (-1)^{e(\Psi(g))} = 0$ and the contribution of these two cancel out. We need to now understand what the fixed graphs are for this involution. The answer is precisely the set of increasing trees.

**Definition** (Increasing Tree)**.** *An increasing tree is a labelled tree, where if we consider any path from the vertex labelled* 1 *to any other vertex, the labels form an increasing sequence.*

**Proposition III.4.1** (Fixed Graphs of the Involution)**.** *The graphs fixed by the involution* $\Psi$ *are precisely increasing trees.*

*Proof.* Firstly we show that if $g$ is an increasing tree, then no edge is $g$-active.

Since $g$ has no cycles, no edge in $g$ is $g$-active. Consider $e = (i, j) \notin E(g)$ and let $k$ be the nearest common ancestor of $i$ and $j$. Let $e' = (k, l)$ be an edge containing $k$ on the path in $g$ from $i$ to $j$. Since $g$ is an increasing tree, $k \leq \min\{i, j\}$ and $l \leq \max\{i, j\}$, so that $e' < e$, so $e$ is not $g$-active.

To show a graph that has no $g$-active edge must be an increasing tree.

$g$ must be a tree, since if $g$ had a cycle, the minimal edge in the cycle is $g$-active. If we have a sequence of labels $1 = i_1 < i_2 < \cdots < i_r > i_{r+1}$ on a path in $g$, starting from vertex $i_1 = 1$, then the edge $(i_{r-1}, i_{r+1})$ is $g$-active. Hence the tree must be increasing. □

Increasing trees on $[n]$ are in bijection with permutations on $[2, n]$. This can be seen through the following mapping from increasing trees to bijections. Take 1 as the root of the tree and for each subtree in the forest formed by removing 1, create a cycle for the permutation. At each stage the parent goes to the oldest sibling. If we are at a leaf, we retrace the steps we made until we come to a branching point (the closest ancestor with multiple children). In this case we take the next largest child and continue like this, when nothing is left we are sent to the root of this subtree.

To see that this is a bijection, we first indicate how to create such a graph from the data of the permutation. Given $\sigma \in \mathcal{S}_{[2,n]}$, we create the tree as follows:

Taking each cycle in turn, we start at the smallest element in the cycle. We attach this element, $r$, to 1. We take $\sigma(r)$ and attach it to $r$. In general, after we have placed $a$, we place $\sigma(a)$ in the following way:

- If $a < \sigma(a)$, then $\sigma(a)$ is attached to $a$ as its parent.
- If $a > \sigma(a)$, then $\sigma(a)$ is attached to the first direct ancestor of $a$, which is $< \sigma(a)$.

- $\sigma(a) = a$ is not possible, since it would just be a single root attached to 1

Hence every permutation may be interpreted as an increasing tree. The mapping is injective, since somewhere on the tree one vertex will have a different largest child and thus correspond to a permutation sending it somewhere else.

In Figure III.3, we see two examples of this relationship.

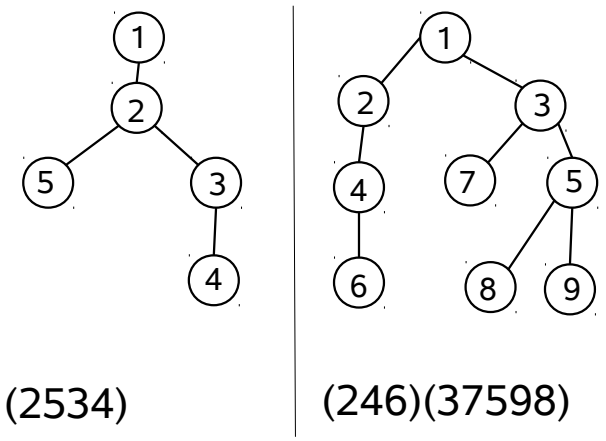



Figure III.3: Two Examples of Increasing Trees and their corresponding Permutations

Hence the number of increasing trees is $(n-1)!$ and they have weight $(-1)^{n-1}$, since each tree as $n-1$ edges. This gives precisely the required $(-1)^{n-1}(n-1)!$ term.

## III.5 The Fixed Graphs for the Two-Connected Case

This section presents the combinatorial identities as theorems and gives an explanation of the fixed graphs, which is my original result.

**Theorem III.5.1** (Combinatorial Identity from the one-particle hard-core model)**.** *We have the combinatorial identity:*

$$\sum_{g \in \mathcal{B}[n]} (-1)^{e(g)} = -(n-2)! \tag{III.5.1}$$

The proof of this identity is given through a different involution $\Psi$, which effectively pairs graphs differing by only one edge, leaving some small collection of graphs fixed, which give the $(n-2)!$ factor. The involution $\Psi$ is defined in Section III.6.

The fixed graphs are described inductively: take the base graph as the single edge between two points with labels 1 and 2. Given a fixed graph on $n$ vertices, we shift its labels by 1 ($i \mapsto i+1$) and add a vertex with label 1. We add from this new vertex an edge to the vertex now labelled 2 and another edge to any of the $n-1$ remaining vertices.

These fixed graphs all have $2n-3$ edges, giving the $-1$ factor and if $a_n$ denotes the number of them, we have the recursion: $a_n = (n-2)a_{n-1}$, which gives the $(n-2)!$, taking $a_2 = 1$.

The fixed graphs for $n = 3, 4$ are shown in Figure III.4.

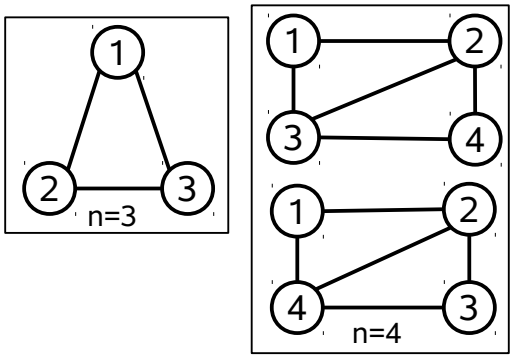



Figure III.4: The Fixed Points for $\Psi$ for $n = 3, 4$

**Theorem III.5.2** (Combinatorial Identity from the continuum hardcore gas)**.** *We have the combinatorial identity:*

$$\sum_{g \in \mathcal{B}[n]} (-1)^{e(g)} \, Vol(\Pi_g) = -n(n-2)! \tag{III.5.2}$$

This theorem is given an alternative proof through a collection of involutions $(\Psi_{\mathbf{h}})_{\mathbf{h} \in \mathbb{Z}^{n-1}}$. The index $\mathbf{h}$ is related to the partition of the polytopes into areas of equal volume attributed to Lass in [Ber08, DLL07]. The meaning of $\mathbf{h}$ is explained in Section III.3.

The fixed points of these involutions occur only when $\mathbf{h}$ is of the form $(0, \cdots, 0, -1, \cdots, -1)$, meaning that any edge is possible. There are precisely $n$ possibilities of these sequences, which corresponds to the $n$ positions of the last zero. The particular $\mathbf{h}$ provides a bijection $\sigma : [n] \to [n]$ on which the fixed graphs correspond to an increasing tree (given by the order $\sigma(i) < \sigma(j)$ if and only if $i < j$) on the labels $\{\sigma(1), \cdots, \sigma(n-1)\}$. This is paired with every edge from $\sigma(n)$ to the vertices $\{\sigma(1), \cdots, \sigma(n-1)\}$.

The number of these increasing trees on $n-1$ vertices is $(n-2)!$ and hence we obtain the factor $n(n-2)!$. We notice that these graphs are on $n-2+n-1 = 2n-3$ edges, which provides the minus sign. An example of a fixed graph is given in Figure III.5.

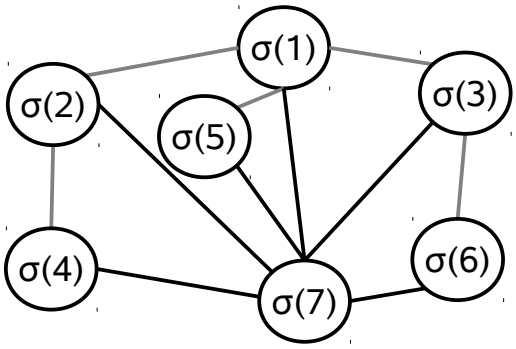



Figure III.5: Example of Fixed Graph for $\Psi_{\mathbf{h}}$ on 7 vertices with the bijection $\sigma$

**Remark 19** (Connection to Bernardi's Interpretation)**.** *It is worth noting that the increasing tree idea present in [Ber08] also appears as an important idea in the fixed points of this involution.*

**Remark 20** (Alternative Version of Theorem III.5.1)**.** *The understanding of the fixed points of the involution $\Psi_{\mathbf{0}}$ provides us with an alternative method of proving Theorem III.5.1. We can see that the fixed points may also correspond to the collection of increasing trees on $[n-1]$ with vertex $n$ having an edge to all other vertices.*

**Remark 21** (Complications for two-connected graphs)**.** *The two-connected case is necessarily more complicated than the connected case. First of all, minimal two-connected graphs do not all have the same number of edges for a fixed number of vertices, although trees (minimal connected graphs) do. Simply removing edges appropriately down to a minimal graph cannot provide a combinatorial understanding as there will still be sign differences to take care of. When we derive the virial expansion for the two models, we find that the virial coefficients all have the same sign. This is present in the combinatorial understanding, since the sign of the $(-1)^{2n-3}$ factor, which appears since the edge-weights are negative, is independent of $n$.*

**Remark 22** (Why to expect $2n-3$ as an important number of edges)**.** *It can be shown that any two-connected graph on $n$ vertices with greater than or equal to $2n-3$ edges, necessarily has a chord. The chord can be removed, keeping the graph*

*two-connected, which indicates that a graph with this number of edges cannot be minimally two-connected. It is also possible to construct a graph with n vertices and* $2n-4$ *edges that is minimally two-connected, as shown in Figure III.6. The number of edges being* $2n-3$ *marks some transition in the possibility of being minimal.*

*Any two-connected graph on n vertices with* $\geq 2n-3$ *edges has a chord.* This is done by induction on the number of vertices $n$.

The cases $n = 2, 3$ are vacuous and one can see from the examples in Figure III.7 that this holds when $n = 4$.

Let $d_1$ be the degree of the vertex labelled 1. If we remove the vertex 1 and its incident edges, then the graph remains connected and can be decomposed into its bc-tree. Each block with $l$ vertices in the tree has to have $\leq 2l-4$ edges or else we have a smaller graph which has a chord by induction. We note here that blocks of size two or three need to be treated separately. We let $l_i$ denote the size of the $i$th block not of size two or three and $B_2$ and $B_3$ denote the number of blocks of size two and three respectively. We know the relationship:

$$\sum_i (l_i - 1) + B_2 + 2B_3 = n - 2 \tag{III.5.3}$$

The total number of edges in the graph must then not exceed:

$$\sum_i 2(l_i - 1) - 2B_{\geq 4} + B_2 + 3B_3 \leq 2n - 4 - B_2 - B_3 - 2B_{\geq 4} \tag{III.5.4}$$

where $b_{\geq 4}$ indicates the number of blocks with more than four vertices. We know that the number of edges must be greater than $2n - 3 - d_1$ and so we obtain the inequality:

$$d_1 \geq 1 + B_2 + B_3 + 2B_{\geq 4} \geq 1 + \text{ total number of blocks} \tag{III.5.5}$$

By the pigeonhole principle, 1 must meet a block at two vertices, call these $\alpha$ and $\beta$. If we have two blocks then we have a third neighbour of 1, call this $\gamma$.

If all three are in the same block then we can find a path $\alpha \to \beta \to \gamma$ and thus the edge $(1, \beta)$ is a chord. Otherwise $\gamma$ is in another block. Let $A$ be the articulation point of the block containing $\alpha$ and $\beta$ closest to $\gamma$. We have a path from $A$ to $\gamma$ outside of this block since it is a connected graph. We are also able to construct a path $\alpha \to \beta \to A$ since they are all in one block. Concatenating these paths gives again a path $\alpha \to \beta \to \gamma$ from which we determine $(1, \beta)$ is a chord.

If we have only one block and two neighbours of one, we may use induc-

tion, since the graph without 1 is therefore two-connected and satisfies the same inequality, since we have taken away two edges.

□

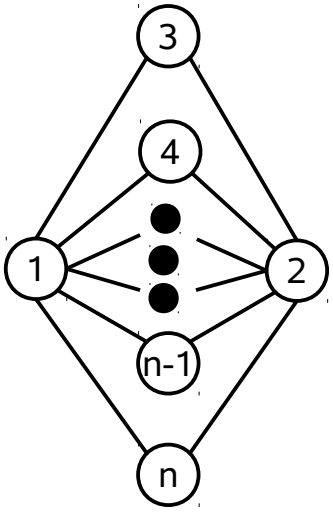



Figure III.6: A graph with $2n - 4$ edges and $n$ vertices which is minimally two-connected

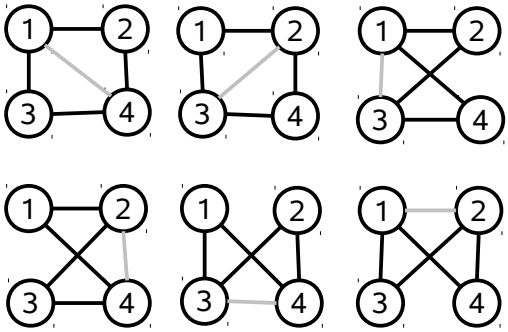



Figure III.7: The chords in graphs of $n = 4$ vertices, excluding the complete graph

## III.6 The Hardcore One Particle Gas - Proof of Theorem III.5.1

The key idea, which is present in the paper by Bernardi [Ber08], is that a graph involution involving removing/adding edges means that the combinatorial factor comes from precisely the number of fixed graphs of the involution. This section describes the involution and proves it does what is required.

In analogy with Bernardi's externally active edges in graphs, we make some definitions here to describe the required involution.

**Definition** (2-active edges)**.** *We call an edge $e$ 2-active for a graph $g$, if the endpoints of $e$ are contained in a cycle comprising of edges $f > e$ in $g$. If a graph $g$ has a 2-active edge, then we define $\varepsilon_g$ to be the maximal such edge.*

For graphs without a 2-active edge, we make a further definition.

**Definition** ($2^\star$-active edges)**.** *We call an edge, $e$, $2^\star$-active for a graph $g$, if the end points of $e$ lie in a cycle in $g$ without the edge $e$.*

*For a graph $g$ with no 2-active edges, we define $A^\star(g)$ as the collection of all $2^\star$-active edges in $g$.*

We use the notation $S^{<e} := \{s \in S | s < e\}$.

**Definition** (Lower active edges)**.** *We call an edge $e$* lower active *for a graph $g$, if it is $2^\star$-active and satisfies:*

$$A^\star(g)^{<e} = A^\star(g \oplus e)^{<e} \tag{III.6.1}$$

**Definition** (Upper active edges)**.** *We call an edge $e$* upper active *for a graph $g$, if it is $2^\star$-active and for all $f \in A^\star(g)^{<e}$, we have:*

$$A^\star((g \oplus e) \oplus f)^{<f} \neq A^\star(g \oplus e)^{<f} \tag{III.6.2}$$

**Definition** (Middle active edges)**.** *We call an edge $e$* middle active *for a graph $g$ if it is both upper and lower active. If it exists, we call the* least *middle active edge for a graph $g$, $\varepsilon_g^\star$.*

The involution $\Psi : \mathcal{B} \to \mathcal{B}$ is described as follows:

i) For a graph $g$ with a 2-active edge, we map $g \mapsto g \oplus \varepsilon_g$.

ii) For a graph $g$ without a 2-active edge, but with a middle active edge, we map $g \mapsto g \oplus \varepsilon_g^\star$.

iii) The remaining graphs are left fixed.

In order to prove that this mapping $\Psi$ does what we require, we break things down into separate lemmata:

- Firstly we show that $\Psi$ is a mapping $\mathcal{B} \to \mathcal{B}$
- We show $\Psi$ is indeed an involution and thus a bijection
- We indicate no graph on $n$ vertices with strictly greater than $2n-3$ edges can be fixed
- We indicate no graph on $n$ vertices with strictly less than $2n-3$ edges can be fixed
- We find the subset of graphs on $n$ vertices with $2n-3$ edges which are fixed

**Lemma III.6.1.** *The image of the mapping $\Psi : \mathcal{B} \to \mathcal{G}$ is contained within $\mathcal{B}$.*

*Proof.* The mapping $\Psi$ can add an edge, remove an edge or leave the graph fixed. If the graph is left fixed, then trivially we get a two-connected graph.
If an edge is added to a two-connected graph, then it remains two-connected.
If an edge $e = (i, j)$ is removed from the graph $G$, the definition of the mapping $\Psi$ ensures that there is a cycle in $G \setminus (i, j)$ containing $i$ and $j$.
$G \setminus (i, j)$ has a cycle including $i$ and $j$ and so they must be in the same block. If we have another block in the graph, then adding the edge $(i, j)$ back in cannot reduce the number of blocks, since it is internal to a block and this would imply $G$ is not two-connected. Therefore, $G \setminus (i, j)$ has to be two-connected.

□

**Lemma III.6.2.** *The mapping $\Psi$ is an involution and thus bijective.*

*Proof.* For 2-active edges, the addition or removal of the largest 2-active edge $\varepsilon_g$ cannot change its status as 2-active, since this depends on larger edges contained in the graph, which are unchanged. Furthermore, it cannot create any larger 2-active edges as these depend on larger edges, which are common between the graphs.

For middle active edges, we note that the upper active condition prevents any smaller edges becoming lower active and so the least middle active edge retains this property under the map $\Psi$. □

**Lemma III.6.3.** *No graph on $n$ vertices and strictly greater than $2n - 3$ edges is fixed.*

*Proof.* It is sufficient to prove that we always have a 2-active edge. In this case, we can always find the greatest such edge and this would give the involution.

This is proved by induction on the number of vertices $n$, for $n \geq 4$.

For $n = 4$, the only such graph we have is the complete graph, which certainly has the 2-active edge $(1, 2)$, illustrated in Figure III.8.

For general $n$, by induction, assume it is true for all $l$ such that $4 \leq l < n$.

Removing the vertex labelled 1 from such a graph, we obtain a connected graph with a corresponding bc-tree. Each of the blocks in this subgraph has (strictly) less than $n$ vertices. If this graph were to not have a 2-active edge, we consider that each block must separately not have a 2-active edge. Because induction starts at $n = 4$, the blocks may be the single edge graph (on two vertices) or the triangle graph (on three vertices), otherwise, the induction hypothesis holds and we have $\leq 2k_i - 3$ edges for every block of size $k_i \geq 4$.

We now indicate how the restrictions on block sizes enforces a 2-active edge to exist with an endpoint at 1.

Counting the vertices in this subgraph, we have:

$$\sum_{\text{blocks of size}\geq 4} (k_i - 1) + 1 = n - 1 - B_2 - 2B_3 \tag{III.6.3}$$

as described in subsection III.2.

If we now count the edges in this subgraph, we have:

$$\begin{aligned} e(G) &\leq \sum_{\text{blocks of size}\geq 4} (2k_i - 3) + B_2 + 3B_3 \\ &= 2 \sum_{\text{blocks of size}\geq 4} (k_i - 1) - B_{\geq 4} + B_2 + B_3 \\ &= 2(n - 2 - B_2 - 2B_3) - B_{\geq 4} + B_2 + 3B_3 \\ &= 2(n-2) - \text{ number of blocks} \end{aligned} \tag{III.6.4}$$

We may also count the total number of edges in this subgraph as $\geq 2n - 2 - d_1$, where $d_1$ is the degree of vertex 1. We thus achieve the inequality:

$$d_1 \geq \text{ number of blocks } + 2 \tag{III.6.5}$$

We know that the number of blocks $\geq 1$ and so $d_1 \geq 3$. By pigeonhole principle, one block has at least two edges from 1. Furthermore, either two blocks have two edges or one block has three.

If a block has three edges from 1, then let $\alpha$ be the label of the minimal such vertex. We prove now that $(1, \alpha)$ is 2-active.

Let $\alpha < \beta < \gamma$ be the labels of these three vertices. We indicate the existence of a path $\beta \to \alpha \to \gamma$ inside this block.

Since we are in a two-connected graph, we have two disjoint paths between any pair of vertices. We may follow a path $\beta \to \alpha$ until it first coincides with any of the two disjoint paths $\alpha \to \gamma$ we then follow this first path to $\alpha$ and then the disjoint path from $\alpha \to \gamma$. This gives a path (with no repeated edges). We may add the two edges $(1, \beta)$ and $(1, \gamma)$ to this to obtain the required cycle. All edges in this cycle are larger than $(1, \alpha)$, since if they involve 1, then the other endpoint is larger (either $\beta$ or $\gamma$), or both endpoints of the edge are larger than 1.

If we have two neighbours of 1 in block $B$ and another two in a disjoint block $C$, we let $\alpha$ be the smallest label of these four and $\beta$ its partner and $\gamma$ and $\delta$ the labels of the other pair. In the bc-tree there is a single path from $B$ to $C$. Let $A_B$

and $A_C$ be the articulation points adjacent to $B$ and $C$ respectively in the path. By above, we can construct paths: $\beta \to \alpha \to A_B$ inside $B$; $A_B \to A_C$ outside of the blocks; and $A_C \to \gamma$ in block $C$. We can therefore add the edges $(1, \beta)$ and $(1, \gamma)$ and we have the required cycle. This then means that $(1, \alpha)$ is 2-active.

□

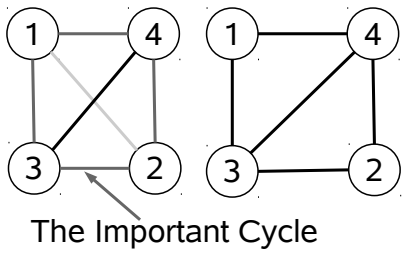



Figure III.8: The Cycle Found in the Graph $K_4$

**Remark 23** (Two connected implies two vertex disjoint paths)**.** *If we do not have the possibility of two disjoint paths between two distinct points, then we have a vertex that must be contained in all paths from these two disjoint vertices. This vertex when removed would then disconnect the two endpoints from each other, this would contradict the graph being two-connected.*

**Corollary III.6.4.** *If $d_1 \geq$ number of blocks $+2$, then we have a 2-active edge and the graph is not fixed.*

**Lemma III.6.5.** *If $d_1 \leq$ number of blocks, then the graph $g$ has either a 2-active edge or a middle active edge.*

*Proof.* As a simple inequality, we know that the number of blocks in a graph is always bounded above by $n - 1$, where $n$ is the number of vertices in the graph. Since $n - 2 \geq$ number of blocks $\geq d_1$, we therefore have a vertex to which 1 is not connected to.

If we have a vertex not in the neighbourhood of 1, which is in a block which is not a leaf block or it is an articulation point between two leaf blocks in the bc-tree, then the smallest such edge is lower active, since adding this edge will not make any further smaller edges a chord. If this edge is not upper active, meaning that if it were added, then there is a smaller edge to remove, which does not change the status of any smaller edge being a chord, then this edge is lower active for the

original graph. The smallest of such edges is also upper active and hence middle active as required.

If all the vertices contained in every block that is not a leaf are neighbours of 1, there is a vertex in a leaf block that is not a neighbour of 1. We may remove the smallest edge $(1, j)$ in one of the internal blocks unless we have a smaller vertex that is not neighbour of 1. This must appear in a leaf block. In this leaf block we have a neighbour of 1, or else the graph with 1 would not be two-connected. If we have a vertex with smaller label than this neighbour of 1, $\nu$, in the leaf vertex, call it $\zeta$, then the edge $(1, \zeta)$ may be added, since we may find a cycle $1 \to \nu \to \zeta \to \eta \to 1$, where $\eta$ is the neighbour of 1 in the closest block to $\zeta$ and is necessarily larger than $\zeta$ or else we would not have such an example.

If all vertices in leaf blocks are larger than the corresponding neighbour of 1 in the block, then none of these edges could be added and removing the smallest edge is fine. □

**Corollary III.6.6.** *In the block decomposition of a graph $g$ without 1, we require the following equality or else we have a 2-active or middle active edge:*

$$d_1 = 1 + \textit{number of blocks} \tag{III.6.6}$$

**Lemma III.6.7.** *In the block decomposition of the graph without 1, we may only have one block.*

*Proof.* From Corollary III.6.6, we have the equality $d_1 = 1+$ number of blocks.

We suppose for contradiction that $d_1 \geq 3$, which is equivalent to assuming we have at least two blocks. We have one block containing two neighbours of 1 and a neighbouring block containing one. We call these vertices $\alpha$ $\beta$ and $\gamma$ and the corresponding articulation point $A$. From the proof of Lemma III.6.3, we can construct in the first block the path $\alpha \to \beta \to A$ and then in the second block $A \to \gamma$, thus giving that $\beta$ is a chord. We therefore have at least one present edge $(1, l) \in G$ with endpoints in a cycle, not including the edge. We call the smallest such $l$, $j$.

This edge is lower active, since no smaller chords are added or removed. If the edge is upper active as well, then the graph is not fixed. We now suppose it is not upper active, then there is some $i < j$ such that $(1, i)$ can be added to $G \setminus \{(1, j)\}$.

Firstly, if $(1, i)$ is in a cycle excluding smaller edges, then it is 2-active and the graph is not fixed.

We know that the only edges involving 1 which therefore may not be removed as a chord are those found on their own in leaf blocks. We note that $i$ is necessarily

smaller than any other neighbour of 1 and that if two leaf blocks are adjacent then that is the whole graph and we would have two neighbours of 1 in one of these blocks, so we may exclude such a possibility, therefore $(1, i)$ has to have all its cycles including at least one of these leaf-block vertices and one such vertex $\mu$ has to appear in all of them. This means that there is an articulation point, in the graph with 1 removed, between $i$ and the neighbours of 1, excluding $\mu$. Hence, $i$ must be in the same block as $\mu$.

If $\mu < i$, then, as before, we can create a path $i \to \mu \to A \to$ the closest vertex attached to 1. This would mean that we cannot add $1 - i$ in the first case and so removing $1 - j$ is fine.

If $\mu > i$, then, since we are assuming no two leaf blocks are adjacent, we have an adjacent block, via articulation point $A$, with a vertex $l > i$, which is a neighbour of 1. The cycle $1 \to \mu \to i \to A \to l \to 1$ is thus made up of larger edges than $(1, i)$ and it could be added in the first place.

Hence a fixed graph cannot have more than one block in the bc-tree without 1 and 1 must have degree two.

□

**Lemma III.6.8.** *No graph with $n$ vertices and strictly less than $2n - 3$ edges is fixed.*

*Proof.* To prove this result, we indicate that we must always have a 2-active or middle active edge in our graph.

We prove this by induction on $n$. We initiate at $n = 4$ and we have only cycles. All these cycles have an edge to add via $\Psi$, as indicated by the dashed line in Figure III.9. Consider a given two-connected graph on $n$ vertices and $\leq 2n - 4$ edges. If we remove 1 and its incident edges, then we are left with a connected graph on the rest of the labels.

We know from Lemma III.6.7, that we have one block and two edges. If we do not have edge $(1, 2)$, then it is 2-active and so we assume we have $(1, 2)$ and one other edge $(1, j)$. Since we have only two edges from 1, this block has less than $2(n - 1) - 3$ edges on $n - 1$ vertices and we may apply the inductive hypothesis to obtain a 2-active edge or a middle active edge for the block. In the former case, it is easy to see that if an edge is 2-active for a subgraph, then it is for the whole graph. We note that a lower active edge in this block retains its lower active property as all edges with 1 as an endpoint are $2^\star$-active, except the two that are present and this is independent of what edges we have in the block. For the upper active condition, we have to check, whether adding an edge from 1 would retain the collection of $2^\star$-

active edges. We realise, independent of edges in the block, adding any edge from 1 would make $(1,2)$ $2^\star$-active and hence we retain the condition for the edge being upper active if the smaller edge is taken to be an edge with endpoint 1. □

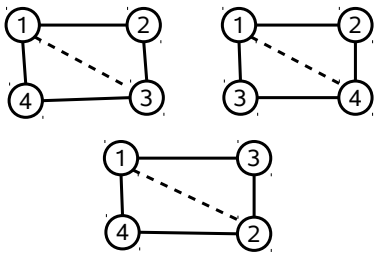



Figure III.9: The Four Examples on $k = 4$ Vertices with Additional Edge

**Lemma III.6.9.** *The fixed points of $\Psi$ on $n$ vertices with $2n-3$ edges are as described in Theorem III.5.1.*

*Proof.* We know from Lemma III.6.7 for a graph to have no 2-active or middle active edge, we have one block with two edges from 1. As explained in the proof of Lemma III.6.8, we must have $(1,2)$ as an edge and an additional edge. If the block is not of the form prescribed, then by induction, argued as in Lemma III.6.8, we have a 2-active or middle active edge for the whole graph. Hence the only possibilities for fixed graphs are those of the prescribed form.

We must now indicate that all graphs of the prescribed form are fixed. We know that such a graph has no chord $e$ that is within a cycle containing edges $f > e$. Any cycle involving the vertices $a$ and $b$ with $a < b$ must contain either an edge $(\alpha, \beta)$ with $\min\{\alpha, \beta\} < a$ or $(a, a+1)$, since the edges from $a$ either go to a smaller vertex, to $a+1$ or to precisely one other larger vertex. Therefore we always have a smaller edge, since if $b > a+1$ then either of the two cases will do. If $b = a+1$, then we couldn't use the edge $(a, a+1)$ if it were to be a chord and we would have a smaller edge.

We have a chord to remove, but for such a chord there is always a smaller chord to add if we remove it. Say $(\alpha, \beta)$ is a chord we are wanting to remove with $\alpha < \beta$. $\alpha$ is only a neighbour of two vertices with a larger label. The cycle involving $\alpha$ and $\beta$ but not the edge $(\alpha, \beta)$, must therefore contain a vertex smaller than $\alpha$ as a neighbour of $\alpha$. Call this vertex $\gamma$. This vertex can have $\beta$ as its other neighbour, in which case, we have a path $\beta \to \alpha$ in this cycle involving a distinct vertex $\delta$, which is not a neighbour of $\gamma$ and so $(\gamma, \delta) < (\alpha, \beta)$ and is a chord in a cycle and

may be added. If $\gamma$ and $\beta$ are not neighbours, then they are in a cycle with each other and the chord $(\gamma, \beta) < (\alpha, \beta)$ and so could have been added. This means that we are unable to remove the chord $(\alpha, \beta)$ as we would have the opportunity to add a smaller chord and hence this graph is fixed.

□

## III.7 The Tonks Gas - Proof of Theorem III.5.2

We define an involution $\Psi_{\mathbf{h}}$ for each $\mathbf{h} \in \mathbb{Z}^{n-1}$ on the set of two connected graphs, which are compatible with the vector $\mathbf{h}$, $\mathcal{B}_{\mathbf{h}}$. In order to do this we must define some order on the edges, which emphasises the compatibility with $\mathbf{h}$. The order depends on $\bar{\mathbf{h}}$ in the sense that we first order edges by the value of $|\bar{h}_i - \bar{h}_j|$, which we call the edge length, and then order within these subsets by the lexicographic order on pairs $(\bar{h}_i, \bar{h}_j)$. The important thing is that for a graph on $n$ vertices, we can understand that this relates to a bijection $\sigma : [n] \to [n]$ in which $\sigma(i)$ indicates the label of the $i$th smallest entry in $\bar{\mathbf{h}}$. We note also that edges with $|\bar{h}_i - \bar{h}_j| > 1$ are forbidden in this set up. The terms larger and smaller used in this section refer to this order on the edges.

Now that the order has been defined, we must indicate what the involution does. This is analogous to the previous case and the fact it is an involution on two-connected graphs follows in the same way. The involution $\Psi_{\mathbf{h}}$ does precisely as $\Psi$, excepting that we have to consider a different order on the edges. The order defines what edges are 2-active, lower upper and middle active for a graph $g$ and we perform the same operation.

In order to prove this result we are required to prove:

- The image of $\mathcal{B}_{\mathbf{h}}$ under $\Psi_{\mathbf{h}}$ is contained in $\mathcal{B}_{\mathbf{h}}$.
- $\Psi_{\mathbf{h}}$ is an involution
- No graphs on $n$ vertices with strictly greater than $2n-3$ edges are fixed
- No graphs on $n$ vertices with strictly less than $2n-3$ edges are fixed
- Those graphs on $n$ vertices and $2n-3$ edges that are fixed are of the prescribed type

The first two requirements follow straightforwardly as in the previous case. The first is just a restriction, noting that no forbidden edges are ever added when using $\Psi_{\mathbf{h}}$. The rules are the same as in the previous case, but with different orderings and so it is an involution.

**Lemma III.7.1.** *No graph on $n$ vertices with strictly greater than $2n-3$ edges is fixed.*

*Proof.* We proceed by induction on $n$. We prove that $g$ has a 2-active edge.

Consider the minimal edge in the above ordering. Either this edge is absent, in which case it is 2-active, since the endpoints will be in a cycle, and we are done, or else it is present. If it is present and its endpoints are in a cycle without the edge (i.e. it is a chord), then it is 2-active and we are done. Therefore, we assume that the minimal edge is present and not a chord. If we remove this edge, we have a connected graph with at least two blocks, since if the two endpoints were in the same block, then we would have a cycle including them both. Each block of size $l$ must have $\leq 2l-3$ edges, or else we can use induction and find a 2-active edge. The total number of edges in this graph is $\geq 2n-3$, since we have removed one edge, and if we count these block by block, we achieve:

$$\begin{aligned} 2n-3 &\leq \sum_{i\in I}(2k_i-3) \\ &= 2\sum_{i\in I}(k_i-1) - \text{number of blocks} \\ &= 2n-2-\text{number of blocks} \end{aligned} \tag{III.7.1}$$

Therefore, $1 \geq$ number of blocks, which gives us a contradiction. Hence we have an edge to add/remove. □

**Lemma III.7.2.** *The only possible fixed graphs under the above involution are those on $2n-3$ edges where $\sigma(n)$ is attached to all other vertices and on the subset of vertices $\{\sigma(1),\cdots,\sigma(n-1)\}$ we have an increasing tree, where the vertex order is defined by $\sigma(i)<\sigma(j)$ if and only if $i<j$. We call the class of graphs $\mathcal{B}_{\mathcal{T}}$ - graphs. The proof also covers the unproved case that no graph on $n$ vertices with $<2n-3$ edges can be fixed by $\Psi$.*

*Proof.* We proceed by induction. For the induction assumption, we have that the only graphs fixed under $\Psi$ are those of the form $\mathcal{B}_{\mathcal{T}}$. For any $4 \leq k < n$. The initiation at 4 can be seen in the appendix, Section III.8.

**[i] The case when** $e(g) = 2n-3$

If we consider the minimal edge of a graph $g$ on $2n-3$ vertices and call it $(\sigma(i),\sigma(j))$. As before, we require that this edge is present in the graph and not a chord, or else it is 2-active. This means that the graph without this edge, $h = g \setminus \{(\sigma(i),\sigma(j))\}$, cannot be two-connected. It must be connected as neither $i$

nor $j$ can be articulation points. Hence we are left with a connected graph with at least two blocks.

Each of the blocks, formed by removing this edge, must have $2l-3$ edges, where $l$ is the number of vertices in the block. If it did not, then we have 2-active or middle active edges from the subgraphs by following the process for blocks with number of edges not equal to $2l-3$. This is by the induction assumption.

Let us have blocks of sizes $(k_i)_{i\in I}$, where $I$ is an index set for the blocks, then we know that $\sum_{i\in I}(k_i-1)=n-1$ and we may add the edges in this graph $H$ in two ways: firstly, by the edges in the blocks:

$$\sum_{i\in I}(2k_i-3)=2\sum_{i\in I}(k_i-1)-|I|=2n-2-\text{number of blocks} \tag{III.7.2}$$

and secondly, from knowing we have just removed an edge from a graph with $2n-3$ edges, we must have $2n-4$ edges. This gives that the number of blocks is two.

These two smaller blocks by induction must be of the form given by $\mathcal{B}_{\mathcal{T}}$, otherwise we can add or remove edges inside the blocks. We want to consider the single articulation point in this connected graph. We wish to first prove that this articulation point is necessarily $\sigma(n)$.

If we suppose for contradiction that the articulation point in the graph is some $a\neq\sigma(n)$. We let $\max_1$ and $\max_2$ be the maximum vertices in the two blocks. We then know by induction that each of these must be connected to all vertices in their respective blocks. At least in one of the two blocks $a\neq\max_i$. Consider the minimal edge adjacent to $a$, which is necessarily in one of the two blocks. If the minimal edge $(a,\mu_1)$ is in the first block and $a\neq\max_1$, then we may consider the cycle: $\max_2\to a\to\max_1\to\mu_1\tilde{\longrightarrow}\sigma(i)\to\sigma(j)\to\max_2$. $\to$ represents a direct edge and $\tilde{\longrightarrow}$ indicates following a path in the increasing tree in the block. If $a=\max_1$, then we do not need the edge $a\to\max_1$ and can replace it with just a single vertex.

Thus we have an edge that may be removed or indeed an even smaller edge that can be removed. If we remove this edge, we need to indicate that no smaller edge can be added. We know we cannot add an edge between the two blocks as this would make $(\sigma(i),\sigma(j))$ removable. A smaller edge then has to involve two vertices in a single block.

Hence, we have a removable edge and so the graph is not fixed. The articulation point must therefore be $\sigma(n)$. We therefore require that $\sigma(n)$ is attached to everything. This implies that we cannot have any impossible edges in our graph. We note that requiring the graph to not have any impossible edges is equivalent to saying that $\mathbf{h}$ has to be of the form $(0,\cdots,0,-1,\cdots,-1)$ and so the smallest edge

is necessarily $(\sigma(1), \sigma(2))$, the first two points to be $-1$, or if only the final entry is $-1$ it would be $n$ and 1 respectively. If all entries are 0, then $\sigma$ is the identity. We have that $\sigma(n)$ has all other vertices in its neighbourhood and that there are increasing trees from $\sigma(1)$ and $\sigma(2)$. If we just add an edge between $\sigma(1)$ and $\sigma(2)$ then we still have an increasing tree on the whole graph and hence the fixed graph has to be of this form.

**[ii] The case** $e(g) < 2n - 3$

We consider removing the minimal edge $(\sigma(i), \sigma(j))$ of the graph $G$ and understand the block-cutpoint decomposition of this graph, which must have at least two blocks. The blocks all have less vertices than the whole graph and so must satisfy the induction assumption - that is be of form $\mathcal{B}_{\mathcal{T}}$ and in particular for a block with $l$ vertices have $2l-3$ edges. If we count the number of edges through the blocks and through the inequality, we obtain:

$$\begin{aligned} e(g) &= \sum_{i \in I} (2k_i - 3) = 2(n-1) - \text{number of blocks} \\ &\leq 2n - 5 \end{aligned}$$

In particular, we have that there are at least three blocks. We therefore know that there are a pair of blocks who do not have the articulation point as $\sigma(n)$. In this case, we may repeat the proof as above, using that we have a path from the articulation points that are not in common between the two blocks to $\sigma(i)$ and $\sigma(j)$ instead of the immediate connection $(\sigma(i), \sigma(j)$. The only difference is whether we may add a smaller edge between two blocks. We cannot add such an edge as the only cycle it can be a chord in involves this minimal edge that we removed. Therefore, because priority is given to edges $e$ appearing as chords in cycles with edges $f > e$, then the edge described would be able to be added or at least the maximal such possibility in the block.

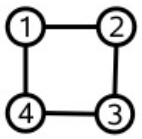
**[iii] Graphs of the form** $\mathcal{B}_{\mathcal{T}}$ **are fixed**

If a graph is of the form described by $\mathcal{B}_{\mathcal{T}}$, then we need to indicate that it is fixed. The only possible edges to add are between non-adjacent vertices in the increasing tree. If we try to add such an edge, then the edges in the increasing tree, which are smaller and are required to form the cycle, mean that it cannot be a chord in a cycle with larger edges. The only edges that may be removed are between $\sigma(n)$ and another vertex. We note that $|\bar{h}_{\sigma(i)} - \bar{h}_{\sigma(j)}| = \frac{|i-j|}{n}$ and so the edges are ordered by the differences of the vertex labels. We note that $\sigma(n-1)$ must be a leaf in this tree and so the edge $(\sigma(n), \sigma(n-1))$ cannot be removed in any case. Therefore, the length of the removable edge is at least $\frac{2}{n}$. We need to indicate that there is a

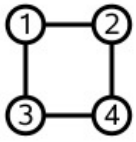
missing edge with a smaller length. If two consecutive labels are not neighbours in the tree then we are done, since they have edge length of $\frac{1}{n}$, otherwise the increasing tree is linear and we can add $(\sigma(1), \sigma(3))$ if we remove any edge from $\sigma(n)$, since it appears earlier in the lexicographic order. □

## III.8 The Examples for $n = 4$

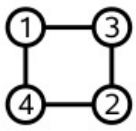
Below are given the tables explaining which graphs are fixed for $n = 4$ for different compatible **h**-values, which are listed beside the graphs. The tables indicate what edges are added/removed according to the rules in Section III.7 and the letters indicate the pairings of the $(G, \mathbf{h})$ pairs through the involution $\Psi_{\mathbf{h}}$.



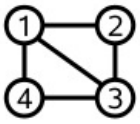

| **h** values | add/remove | $\bar{\mathbf{h}}$ | |
|---|---|---|---|
| $(0,0,0)$ | $(1,3)$ | $(0,\frac{1}{4},\frac{1}{2},\frac{3}{4})$ | A |
| $(0,0,-1)$ | $(2,4)$ | $(0,\frac{1}{4},\frac{1}{2},-\frac{1}{4})$ | C |
| $(0,-1,-1)$ | $(1,3)$ | $(0,\frac{1}{4},-\frac{1}{2},-\frac{1}{4})$ | B |
| $(-1,-1,-1)$ | $(2,4)$ | $(0,-\frac{3}{4},-\frac{1}{2},-\frac{1}{4})$ | D |



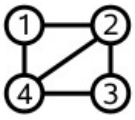

| **h** values | add/remove | $\bar{\mathbf{h}}$ | |
|---|---|---|---|
| $(0,0,0)$ | $(2,3)$ | $(0,\frac{1}{4},\frac{1}{2},\frac{3}{4})$ | H |
| $(0,0,-1)$ | $(1,4)$ | $(0,\frac{1}{4},\frac{1}{2},-\frac{1}{4})$ | E |
| $(0,-1,-1)$ | $(1,4)$ | $(0,\frac{1}{4},-\frac{1}{2},-\frac{3}{4})$ | F |
| $(-1,0,-1)$ | $(1,4)$ | $(0,-\frac{3}{4},\frac{1}{2},-\frac{1}{4})$ | G |
| $(-1,-1,-1)$ | $(2,3)$ | $(0,-\frac{3}{4},-\frac{1}{2},-\frac{1}{4})$ | I |
| $(-1,-1,-2)$ | $(2,3)$ | $(0,-\frac{3}{4},-\frac{1}{2},-\frac{5}{4})$ | J |



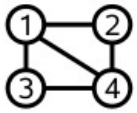

| **h** values | add/remove | $\bar{\mathbf{h}}$ | |
|---|---|---|---|
| $(0,0,0)$ | $(1,2)$ | $(0,\frac{1}{4},\frac{1}{2},\frac{3}{4})$ | K |
| $(0,0,-1)$ | $(1,2)$ | $(0,\frac{1}{4},\frac{1}{2},-\frac{1}{4})$ | L |
| $(0,-1,0)$ | $(1,2)$ | $(0,\frac{1}{4},-\frac{1}{2},\frac{3}{4})$ | M |
| $(0,-1,-1)$ | $(3,4)$ | $(0,\frac{1}{4},-\frac{1}{2},-\frac{3}{4})$ | N |
| $(-1,-1,-1)$ | $(3,4)$ | $(0,-\frac{3}{4},-\frac{1}{2},-\frac{1}{4})$ | P |
| $(1,0,0)$ | $(3,4)$ | $(0,\frac{5}{4},\frac{1}{2},\frac{3}{4})$ | O |



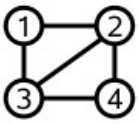

| $\mathbf{h}$ values | add/remove | $\bar{\mathbf{h}}$ | |
|---|---|---|---|
| $(0,0,0)$ | $(1,3)$ | $(0,\frac{1}{4},\frac{1}{2},\frac{3}{4})$ | A |
| $(0,0,-1)$ | FIXED | $(0,\frac{1}{4},\frac{1}{2},-\frac{1}{4})$ | |
| $(0,-1,-1)$ | $(1,3)$ | $(0,\frac{1}{4},-\frac{1}{2},-\frac{1}{4})$ | B |
| $(-1,-1,-1)$ | FIXED | $(0,-\frac{3}{4},-\frac{1}{2},-\frac{1}{4})$ | |



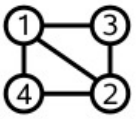

| $\mathbf{h}$ values | add/remove | $\bar{\mathbf{h}}$ | |
|---|---|---|---|
| $(0,0,0)$ | FIXED | $(0,\frac{1}{4},\frac{1}{2},\frac{3}{4})$ | |
| $(0,0,-1)$ | $(2,4)$ | $(0,\frac{1}{4},\frac{1}{2},-\frac{1}{4})$ | C |
| $(0,-1,-1)$ | FIXED | $(0,\frac{1}{4},-\frac{1}{2},-\frac{1}{4})$ | |
| $(-1,-1,-1)$ | $(2,4)$ | $(0,-\frac{3}{4},-\frac{1}{2},-\frac{1}{4})$ | D |



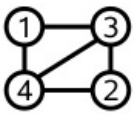

| $\mathbf{h}$ values | add/remove | $\bar{\mathbf{h}}$ | |
|---|---|---|---|
| $(0,0,0)$ | FIXED | $(0,\frac{1}{4},\frac{1}{2},\frac{3}{4})$ | |
| $(0,0,-1)$ | $(1,4)$ | $(0,\frac{1}{4},\frac{1}{2},-\frac{1}{4})$ | E |
| $(0,-1,-1)$ | $(1,4)$ | $(0,\frac{1}{4},-\frac{1}{2},-\frac{3}{4})$ | F |
| $(-1,0,-1)$ | $(1,4)$ | $(0,-\frac{3}{4},\frac{1}{2},-\frac{1}{4})$ | G |
| $(-1,-1,-1)$ | $(2,3)$ | $(0,-\frac{3}{4},-\frac{1}{2},-\frac{1}{4})$ | $\delta$ |



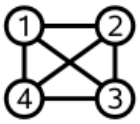

| $\mathbf{h}$ values | add/remove | $\bar{\mathbf{h}}$ | |
|---|---|---|---|
| $(0,0,0)$ | $(2,3)$ | $(0,\frac{1}{4},\frac{1}{2},\frac{3}{4})$ | H |
| $(0,0,-1)$ | $(1,4)$ | $(0,\frac{1}{4},\frac{1}{2},-\frac{1}{4})$ | $\beta$ |
| $(0,-1,-1)$ | FIXED | $(0,\frac{1}{4},-\frac{1}{2},-\frac{3}{4})$ | |
| $(-1,-1,-1)$ | $(2,3)$ | $(0,-\frac{3}{4},-\frac{1}{2},-\frac{1}{4})$ | I |
| $(-1,-1,-2)$ | $(2,3)$ | $(0,-\frac{3}{4},-\frac{1}{2},-\frac{5}{4})$ | J |




| $\mathbf{h}$ values | add/remove | $\bar{\mathbf{h}}$ | |
|---|---|---|---|
| $(0,0,0)$ | $(1,2)$ | $(0,\frac{1}{4},\frac{1}{2},\frac{3}{4})$ | K |
| $(0,0,-1)$ | $(1,2)$ | $(0,\frac{1}{4},\frac{1}{2},-\frac{1}{4})$ | L |
| $(0,-1,0)$ | $(1,2)$ | $(0,\frac{1}{4},-\frac{1}{2},\frac{3}{4})$ | M |
| $(0,-1,-1)$ | $(3,4)$ | $(0,\frac{1}{4},-\frac{1}{2},-\frac{3}{4})$ | $\gamma$ |
| $(-1,-1,-1)$ | FIXED | $(0,-\frac{3}{4},-\frac{1}{2},-\frac{1}{4})$ | |




| $\mathbf{h}$ values | add/remove | $\bar{\mathbf{h}}$ | |
|---|---|---|---|
| $(0,0,0)$ | $(1,2)$ | $(0,\frac{1}{4},\frac{1}{2},\frac{3}{4})$ | $\alpha$ |
| $(0,0,-1)$ | FIXED | $(0,\frac{1}{4},\frac{1}{2},-\frac{1}{4})$ | |
| $(0,-1,-1)$ | $(3,4)$ | $(0,\frac{1}{4},-\frac{1}{2},-\frac{3}{4})$ | N |
| $(-1,-1,-1)$ | $(3,4)$ | $(0,-\frac{3}{4},-\frac{1}{2},-\frac{1}{4})$ | P |
| $(1,0,0)$ | $(3,4)$ | $(0,\frac{5}{4},\frac{1}{2},\frac{3}{4})$ | O |




| $\mathbf{h}$ values | add/remove | $\bar{\mathbf{h}}$ | |
|---|---|---|---|
| $(0,0,0)$ | $(1,2)$ | $(0,\frac{1}{4},\frac{1}{2},\frac{3}{4})$ | $\alpha$ |
| $(0,0,-1)$ | $(1,4)$ | $(0,\frac{1}{4},\frac{1}{2},-\frac{1}{4})$ | $\beta$ |
| $(0,-1,-1)$ | $(3,4)$ | $(0,\frac{1}{4},-\frac{1}{2},-\frac{1}{4})$ | $\gamma$ |
| $(-1,-1,-1)$ | $(2,3)$ | $(0,-\frac{3}{4},-\frac{1}{2},-\frac{1}{4})$ | $\delta$ |

## III.9 Outlook and Conclusions

The cancellations in two simple models of the virial expansion can be understood combinatorially, by understanding the involutions given in this paper. It would be helpful to understand how the particular graphs, which are fixed points, can help to give an understanding of better bounds for the virial expansion. The parallel that is useful to draw here is that for the cluster expansion, we have the increasing and rooted trees as the combinatorial objects representing the two cases above. It has been shown by Groeneveld [Gro62] that these examples provide the extreme cases for positive potentials and an adaptation is available for stable potentials.

Much work has been done on the expansions or graph-tree identities in this case in the paper by Bernardi [Ber08] and its references to matroids in the q-state Potts Model [Sok05]. Edges in graphs can be viewed as externally active and these are the edges that are able to be identified in particular factors. The combinatorial part of the argument then relies on counting trees. This has already been understood as the Penrose partition [Pen67]. Is it then possible that we can use internally and externally active edges to find a similar conclusion for positive potentials with these increasing trees and the special vertex adjacent to everything? This should certainly be the next stage of investigation and the the importance of a combinatorial understanding of these coefficients. The issues are that the matroids do not easily generalise into the set of two-connected graphs. It indicates the necessity of having to generalise the key elements of the structure present in the paper to be able to capture two-connected graphs. This is explored further in Chapter VII

# Chapter IV

# Virial Expansion Bounds

This chapter presents bounds on the virial expansion coefficients and the radius of convergence. The simple bound found by Ruelle in [Rue69] is useful and gives the main idea of using complex analysis to obtain bounds. Section IV.1 gives the bounds made by Groeneveld described in [McCoy10], which are the strongest known for positive potentials. The other sections in this chapter are adapted from the paper [Tate13] and they explain a refinement of the Lebowitz Penrose [LePen64] bound and provide comparisons of current bounds. Furthermore, the cluster expansion bounds provided by Procacci [Pro07] and Poghosyan and Ueltschi [PoUe09] are used in this situation to provide slightly different bounds. The material is presented in a general framework, which allows it to be specialised to other cluster expansion bounds.

This chapter also concludes the situation for the virial expansions, indicating areas in which to look for further improvements of bounds, which are featured in the subsequent chapters.

## IV.1 Groeneveld's Bounds

In [McCoy10], McCoy presents Groeneveld's result for positive potentials as:

$$|c_k| \leq \frac{1}{k} a_{k-1} (|2b_2|)^k \tag{IV.1.1}$$

where $c_k$ is the $k$th virial coefficient and $b_2$ is the second Mayer coefficient in the cluster expansion. The expression for $a_k$ is:

$$a_k = \frac{1}{2\pi i} \oint_C \frac{\mathrm{d}\xi}{\xi^{k+1}} e^{\xi} (1 + e^{\xi})^{k-1} \tag{IV.1.2}$$

If we wish to achieve a bound on $|a_k|$, then we can evaluate the integral expression along a contour comprising of a circle of radius $R$ and then minimise the function of $R$ we obtain. Evaluating the contour integral we obtain:

$$|a_k| \leq \frac{1}{R^k} e^R (1+e^R)^{k-1} \tag{IV.1.3}$$

We seek to minimise the function:

$$f(R) = \frac{1}{R^k} e^R (1+e^R)^{k-1} \tag{IV.1.4}$$

Taking the derivative we obtain:

$$f'(R) = \frac{(1+e^R)^{k-2} e^R}{R^{k+1}} (-k(1+e^R) + R(1+e^R) + R(k-1)e^R) \tag{IV.1.5}$$

The prefactor is always (strictly) positive, so to find a turning point, we need to find a zero of the second factor, which gives the equation:

$$(R-1)e^R = 1 - \frac{R}{k} \tag{IV.1.6}$$

for large $k$ the minimum occurs close to the point where $(R-1)e^R = 1$, so we take this value uniformly and get:

$$|a_k| \leq \frac{1}{r^k} \frac{1}{R-1} \left(1 + \frac{1}{R-1}\right)^{k-1} = \frac{1}{R} \frac{1}{(R-1)^k} \tag{IV.1.7}$$

Letting $c_0 = R - 1$, we then obtain:

$$|a_k| \leq \frac{1}{1+c_0} \frac{1}{(c_0)^k} \tag{IV.1.8}$$

where $c_0 e^{c_0+1} = 1$, alternatively in the notation of Lambert W-functions $c_0 = W(e^{-1})$ Hence we have the bound:

$$|c_k| \leq \frac{1}{k} \frac{W(e^{-1})}{1+W(e^{-1})} \left(\frac{|2b_2|}{W(e^{-1})}\right)^k \tag{IV.1.9}$$

Giving the radius of convergence of the virial expansion being:

$$\mathcal{R}_{\text{vir}} \geq \frac{W(e^{-1})}{|2b_2|} \tag{IV.1.10}$$

## IV.2 Virial Expansion Bounds from Cluster Expansion Bounds

Starting with the grand canonical partition function from equation (I.1.3):

$$\Xi(z) = \sum_{N=0}^{\infty} Z_N z^N \tag{IV.2.1}$$

With the configuration integral:

$$Z_N = \frac{1}{N!} \int_{\mathbb{R}^d} e^{-\beta \sum\limits_{1 \le i < j \le N} \Phi(x_i - x_j)} \mathrm{d}^d x \tag{IV.2.2}$$

We obtain the cluster expansion for pressure as:

$$\beta P = \sum_{n \ge 1} b_n z^n \tag{IV.2.3}$$

Using the formula $\rho = z\frac{\partial}{\partial z}(\beta P)$, we have the following corresponding fugacity expansion for density:

$$\rho = \sum_{n \ge 1} n b_n z^n \tag{IV.2.4}$$

We see that the power series expansion for $\rho$ has a zero constant term and non zero $z$ term, where we are assuming that $b_1 \neq 0$ and usually we set it to 1. This means that it is possible to invert the $\rho - z$ relationship, in order to obtain an expansion for $z(\rho)$. This can be substituted into (IV.2.3) in order to get a power series for $P$ in terms of $\rho$:

$$\beta P = \sum_{n \ge 1} c_n \rho^n \tag{IV.2.5}$$

Lagrange inversion techniques are the usual approach to obtaining the coefficients and this has a nice representation as a contour integral. If we want to compute the coefficient $c_n$, we can take the contour integral (around 0) of $\frac{\partial \beta P}{\partial \rho}$ divided by $n\rho^n$. This gives us the formula:

$$c_n = \frac{1}{2\pi i} \oint_C \frac{\frac{\partial \beta P}{\partial \rho}}{n\rho^n} \mathrm{d}\rho \tag{IV.2.6}$$

We can manipulate this equation to get it in terms of the $z$-variable, since we know

about the cluster expansion already.

$$\begin{aligned} c_n &= \frac{1}{2\pi i}\oint_{C'} \frac{\frac{\partial \beta P}{\partial z}}{n\rho^n}\mathrm{d}z \\ &= \frac{1}{2n\pi i}\oint_{C'} \frac{z\frac{\partial \beta P}{\partial z}}{z\rho^n}\mathrm{d}z \\ &= \frac{1}{2n\pi i}\oint_{C'} \frac{\mathrm{d}z}{z\rho^{n-1}} \end{aligned} \tag{IV.2.7}$$

This (implicit) relationship between the virial and the cluster coefficients is our starting point. The idea is that we can bound such an integral quite easily when we consider bounds on $|\rho|$, when we write it in terms of the fugacity $z$ as in (IV.2.4).

The contour $C'$ is the image of the contour $C$, under the mapping from $\rho$ space to $z$ space. The contours are taken so that $C'$ lies within the radius of convergence for the cluster expansion.

This method is what is used by Ruelle [Rue69] and Lebowitz and Penrose [LePen64]. The technique detailed in those papers can be generalised and slightly improved, using the Lambert W-function.

## IV.3 Main Results

The results of this chapter are in the model of the classical gas, as explained above, with the assumptions for the cluster expansion being:

**Assumption 6** (Potential)**.** *The $N$-particle interaction potential: $\mathcal{U}_N(x_1, \cdots, x_N)$ may be written as the sum of pair-potentials:*

$$\mathcal{U}_N(x_1, \cdots, x_N) = \sum_{1\le i<j\le N} \Phi(x_i, x_j) \tag{IV.3.1}$$

*Furthermore, we assume that the pair potentials $\Phi(x_i, x_j)$ are central, that is, they only depend on the distance from $x_i$ to $x_j$.*

**Assumption 7** (Stability)**.** *The potential energy is assumed to be stable, that is, there is a $B > 0$, such that for every $N$ and $(x_1, \cdots, x_N) \in \mathbb{R}^{Nd}$, we have:*

$$\mathcal{U}_N(x_1, \cdots, x_N) = \sum_{1\le i<j\le N} \Phi(x_i, x_j) \ge -BN \tag{IV.3.2}$$

*where $d$ is the dimension of our system.*

**Definition** ($C(\beta)$ and $R(\beta)$ - 'temperedness'). *We have two main functions of $\beta$, which play an important rôle in the cluster expansion bounds:*

$$C(\beta) := \int_{\mathbb{R}^d} \left| e^{-\beta\Phi(0,x)} - 1 \right| \mathrm{d}^d x \tag{IV.3.3}$$

$$R(\beta) := \left( |\mathbb{B}| r^d + \beta \int_{|y|>r} |\Phi(0,y)| \mathrm{d}^d y \right) \tag{IV.3.4}$$

*If the expression (IV.3.3) is finite then the potential $\Phi$ is called 'tempered', which we assume for bounds involving $C(\beta)$*

*The $r$ in (IV.3.4) represents the radius of the hard-core interaction. $|\mathbb{B}|$ is the surface area of a $d$-dimensional sphere.*

We write $\mathcal{R}_{\mathrm{Vir}}$ for the radius of convergence of the virial expansion.

**Definition** (Lambert W-function). *We denote by $W(z)$, the Lambert W-function with domain $\mathbb{R}_+ := \{x \in \mathbb{R} | x \geq 0\}$ and range $\mathbb{R}_+$. It is the solution to:*

$$W(z) e^{W(z)} = z \tag{IV.3.5}$$

**Theorem IV.3.1** (General Virial Bounds). *Assuming cluster coefficient bounds of the form:*

$$|n b_n| \leq a \frac{n^{n-1}}{n!} b^n \tag{IV.3.6}$$

*where $a$ and $b$ are non-negative functions of inverse temperature $\beta$, we have the virial coefficient bounds:*

$$|c_n| \leq \frac{\beta^{-1}}{n} \left( a^{-1} \frac{W\left(\frac{eab}{1+ab}\right)}{\left(W\left(\frac{eab}{1+ab}\right) - 1\right)^2} \right)^{n-1} \tag{IV.3.7}$$

*which gives the lower bound on the radius of convergence as:*

$$\mathcal{R}_{Vir} \geq a \frac{\left(W\left(\frac{eab}{1+ab}\right) - 1\right)^2}{W\left(\frac{eab}{1+ab}\right)} \tag{IV.3.8}$$

If we apply this general theorem, which is derived in Section IV.4, to two specific bounds we have for cluster expansions, we achieve:

**Corollary IV.3.2** (Improved Lebowitz-Penrose)**.** *The cluster expansion bounds:*

$$|nb_n| \leq \frac{n^{n-1}}{n!} e^{2\beta B(n-1)} C(\beta)^{n-1} \tag{IV.3.9}$$

*give the bound for virial coefficients:*

$$|c_n| \leq \frac{\beta^{-1}}{n} \left( C(\beta) e^{4\beta B} \frac{W\left(\frac{e}{1+e^{2\beta B}}\right)}{\left(W\left(\frac{e}{1+e^{2\beta B}}\right)-1\right)^2} \right)^{n-1} \tag{IV.3.10}$$

*and the bound for the radius of convergence:*

$$\mathcal{R}_{Vir} \geq C(\beta)^{-1} e^{-4\beta B} \frac{\left(W\left(\frac{e}{1+e^{2\beta B}}\right)-1\right)^2}{W\left(\frac{e}{1+e^{2\beta B}}\right)} \tag{IV.3.11}$$

This is shown in Section IV.5

For purely hard-core interactions $B = 0$ and the radius of convergence satisfies:

$$\mathcal{R}_{\text{Vir}} \geq C(\beta)^{-1} \frac{\left(W\left(\frac{e}{2}\right)-1\right)^2}{W\left(\frac{e}{2}\right)} \tag{IV.3.12}$$

This is precisely the same as what is obtained by Lebowitz-Penrose.

**Corollary IV.3.3** (Alternative Bounds)**.** *For cluster expansion bounds:*

$$|nb_n| \leq \frac{n^{n-1}}{n!} R(\beta)^{n-1} e^{n\beta B} \tag{IV.3.13}$$

*we have the bound for virial coefficients as:*

$$|c_n| \leq \frac{\beta^{-1}}{n} R(\beta)^{n-1} \left( \frac{W\left(\frac{e^{\beta B+1}}{1+e^{\beta B}}\right)}{\left(W\left(\frac{e^{\beta B+1}}{1+e^{\beta B}}\right)-1\right)^2} \right)^{n-1} \tag{IV.3.14}$$

*Thus giving the radius of convergence as:*

$$\mathcal{R}_{Vir} \geq R(\beta)^{-1} \frac{\left(W\left(\frac{e^{\beta B+1}}{1+e^{\beta B}}\right)-1\right)^2}{W\left(\frac{e^{\beta B+1}}{1+e^{\beta B}}\right)} \tag{IV.3.15}$$

**Remark 24.** *The bounds for the cluster coefficients used in Corollary IV.3.2 are obtained in [Rue69] and [LePen64]. The bounds for the cluster coefficients used in Corollary IV.3.3 are obtained in [PoUe09].*

Having a new separate bound is only a good idea, if it is an improvement in certain cases. In Section IV.5.3 we understand how the two different bounds compare and conclude:

**Proposition IV.3.4** (Comparison of Bounds)**.** *The bound in Corollary IV.3.3 is better than that in Corollary IV.3.2, precisely when*

$$1.6R(\beta) < C(\beta) \tag{IV.3.16}$$

This is likely to happen at low temperatures (large $\beta$) and for potentials with a hard core and attractive tail. The attractive element of the potential is important, since the improvements are made corresponding to the stability parameter $B$.

## IV.4 General Derivation

We wish to obtain bounds on virial coefficients using those which come from cluster coefficients, in order to gain an estimate of the lower bound for the radius of convergence of the virial series. This method loosely follows the method used by Ruelle in [Rue69] and that of Lebowitz and Penrose in [LePen64]. Our starting point is the cluster expansion bound, which is given in the form:

$$|nb_n| \leq a\frac{n^{n-1}}{n!}b^n \tag{IV.4.1}$$

Where $a$ and $b$ are positive functions of inverse temperature $\beta$, defined by the particular bound we use.

We start from the identity (IV.2.7) for the virial coefficients:

$$c_n = \frac{1}{2\pi i}\oint_C \frac{\mathrm{d}z}{nz\rho^{n-1}} \tag{IV.4.2}$$

In order to get an upper bound on $|c_n|$, we need to bound $|\rho|$ from below. To do this we use the bound:

$$|\rho - z| \leq \sum_{n=2}^{\infty} |nb_n||z|^n \tag{IV.4.3}$$

We can then substitute in for the upper bounds we have on the cluster coefficients:

$$|\rho - z| \leq a\sum_{n=2}^{\infty} \frac{n^{n-1}}{n!}(b|z|)^n \tag{IV.4.4}$$

We define a function: $f(x) := \sum_{n=1}^{\infty} \frac{n^{n-1}}{n!} x^n$ and write (IV.4.4) conveniently as:

$$|\rho - z| \leq a(f(b|z|) - b|z|) \tag{IV.4.5}$$

We then use the reverse triangle inequality to obtain:

$$|\rho| \geq |z|(1 + ab) - af(b|z|) \tag{IV.4.6}$$

Let $b|z| = se^{-s}$ and observe that, from the assumed generic bound on the cluster expansion, $b|z| \leq e^{-1}$ in order for $\sum_{n=1}^{\infty} nb_n z^n$ to converge. $se^{-s}$ is an increasing function on $(0, 1)$ and takes values in $(0, e^{-1})$ as required. We then have the bound in terms of $s$:

$$|\rho| \geq b^{-1} se^{-s}(1 + ab) - af(se^{-s}) \tag{IV.4.7}$$

We have chosen the function $se^{-s}$ since it is the inverse function of $f$. This can be understood from Lagrange inversion (Remark 26) and is a result in [CJK97] . We thus have the expression:

$$|\rho| \geq b^{-1} se^{-s}(1 + ab) - as \tag{IV.4.8}$$

We thus seek to maximise the right hand side to get the best possible bound. We notice that in the range $s \in (0, 1)$ we have a zero at $s = 0$ and another when $b^{-1}e^{-s}(1 + ab) - a = 0$ i.e. when $e^s = 1 + \frac{1}{ab}$, so at $s = \ln\left(1 + \frac{1}{ab}\right)$. It is positive for $s \in (0, \ln\left(1 + \frac{1}{ab}\right))$. We seek the value of $s$ to maximise this function in this range.

**Remark 25.** *This approach to estimating a lower bound for the radius of convergence for the virial expansion, takes a value of $|z|$ some distance away from its maximal value for convergent cluster expansions. We expect density to increase with fugacity and so the maximal density for which the cluster expansions remain convergent appears to be greater than that for the virial expansion. This is not true in general and is a weakness of this approach.*

If we define:

$$r(s) := s(e^{-s}(1 + ab)b^{-1} - a) \tag{IV.4.9}$$

and take the derivative to search for an extremum in this range.

$$r'(s) = (e^{-s}(1 + ab)b^{-1} - a) - se^{-s}(1 + ab)b^{-1} \tag{IV.4.10}$$

if we find when $r'(s) = 0$ in (IV.4.10), then we solve:

$$(1-s)e^{-s} = \frac{ab}{1+ab} \tag{IV.4.11}$$

Substituting $\gamma = 1 - s$, we get the equation for $\gamma$:

$$\gamma e^{\gamma} = \frac{eab}{1+ab} \tag{IV.4.12}$$

The Lambert $W$-function, as is explained in [CJK97], is the inverse of $\gamma e^{\gamma}$ and so we can write (IV.4.9) in terms of $\gamma$ and substitute $\gamma$ for $W(\mu)$, where $\mu := \frac{eab}{1+ab}$. For further applications of the Lambert W-function, the paper of Caillol [Cai03] contains good examples. We thus get:

$$\begin{aligned}
\tilde{r}(\gamma) &= (1-\gamma)(e^{\gamma}\frac{1+ab}{eb} - a) \\
&= a(e^{\gamma}\frac{1+ab}{eab} - \gamma e^{\gamma}\frac{1+ab}{eab} - 1 + \gamma) \\
&= a(\frac{1}{\gamma} - 2 + \gamma) \\
&= a\frac{(W(\mu)-1)^2}{W(\mu)}
\end{aligned} \tag{IV.4.13}$$

Where we use (IV.4.12) to cancel: $\gamma e^{\gamma}\frac{1+ab}{eab} = 1$ and $e^{\gamma}\frac{1+ab}{eab} = \frac{1}{\gamma}$.

We evaluate the integral (IV.4.2) along the contour described by the circle $|z|$ =constant, where the constant is determined by the manipulations above. This leaves us with the integral

$$|c_n| \leq \frac{1}{2n\pi}\oint_C \frac{\mathrm{d}z}{|z||\rho|^{n-1}} \tag{IV.4.14}$$

This gives us bounds on the coefficients $c_n$ as:

$$|c_n| \leq \frac{1}{n}\left(a^{-1}\frac{W(\mu)}{(W(\mu)-1)^2}\right)^{n-1} \tag{IV.4.15}$$

This gives us the radius of convergence of the virial expansion, satisfying:

$$\mathcal{R}_{\text{vir}} \geq a\frac{(W(\mu)-1)^2}{W(\mu)} \tag{IV.4.16}$$

**Remark 26** (Lagrange Inversion - Simple Example)**.** *To find the inverse of the function $se^{-s}$ as a power series, we can use the one-dimensional Lagrange inversion form from [MSV06]. We want to find a power series of $s$ in terms of $y = se^{-s}$, which*

*amounts to finding the inverse. We write the equation in the convenient form:*

$$s = ye^s \tag{IV.4.17}$$

*The function $\phi$ on which we wish to perform Lagrange inversion (see Section IX.3) is $\phi(s) = e^s$ and the formula for the nth coefficient in the desired power series is:*

$$\begin{aligned}[y^n]s &= \frac{1}{n}[s^{n-1}]e^{ns} \\ &= \frac{1}{n}\frac{n^{n-1}}{(n-1)!} = \frac{n^{n-1}}{n!}\end{aligned} \tag{IV.4.18}$$

*The power series we obtain is $s = \sum_{n\geq 1} \frac{n^{n-1}}{n!} y^n$ as required.*

**Remark 27** (The Lambert W-function)**.** *The inverse of the function $se^s$ is many valued, but has only one real branch for positive $s$, which is what we use. Its power series around 0 is $-f(-s)$, where $f$ is defined as above. Of course this only gives us an inverse for $|s| < e^{-1}$. The real branch of the inverse, however is well defined for positive $s$ and so it is fine to use it here.*

## IV.5 Relationship of the General Derivation to Previous Bounds

We have the bound on cluster expansions from [Rue69] as:

$$|nb_n| \leq \frac{n^{n-1}}{n!} e^{2\beta B(n-2)} C(\beta)^{n-1} \tag{IV.5.1}$$

Where, the parameter $C(\beta) < \infty$ defines a tempered potential:

$$C(\beta) = \int_{\mathbb{R}^D} \left| e^{-\beta\Phi(\mathbf{x})} - 1 \right| \mathrm{d}^D\mathbf{x} \tag{IV.5.2}$$

A stable potential is also assumed. That is our potential can be written as the sum of pair potentials and there exists a $B > 0$, such that for any $n \in \mathbb{N}$ and any $\mathbf{x}_1, \cdots \mathbf{x}_n \in \mathbb{R}^D$, we have that:

$$\sum_{i\neq j} \Phi(\mathbf{x}_i - \mathbf{x}_j) \geq -Bn \tag{IV.5.3}$$

This corresponds to setting my parameters in the previous section to:

$$a = C(\beta)^{-1}e^{-4\beta B} \tag{IV.5.4}$$

$$b = e^{2\beta B}C(\beta) \tag{IV.5.5}$$

$$ab = e^{-2\beta B} \tag{IV.5.6}$$

$$\mu = \frac{eab}{1+ab} = \frac{e}{1+e^{2\beta B}} \tag{IV.5.7}$$

This gives:

$$|c_n| \leq \frac{1}{n}\left(C(\beta)e^{4\beta B}\frac{W\left(\frac{e}{1+e^{2\beta B}}\right)}{(W\left(\frac{e}{1+e^{2\beta B}}\right)-1)^2}\right)^{n-1} \tag{IV.5.8}$$

and the radius of convergence satisfying:

$$\mathcal{R}_{\text{Vir}} \geq C(\beta)^{-1}e^{-4\beta B}\frac{(W\left(\frac{e}{1+e^{2\beta B}}\right)-1)^2}{W\left(\frac{e}{1+e^{2\beta B}}\right)} \tag{IV.5.9}$$

### IV.5.1 Morais-Procacci Bound

In the recent paper [MoPr13], Morais and Procacci obtain a bound for the virial coefficients via the Canonical Ensemble, in the form of a polymer expansion. The cluster expansion bounds are:

$$|nb_n| \leq \frac{n^{n-1}}{n!}e^{2\beta B(n-2)}C(\beta)^{-1} \tag{IV.5.10}$$

Presented here is an alternative derivation of the Morais Procacci bounds using the method outlined in Section IV.4.

The values $a = C(\beta)^{-1}e^{-4\beta B}$; $b = e^{2\beta B}C\beta$; and thus $ab = e^{-2\beta B}$ are substituted into (IV.4.9), so that we get:

$$r(s) = s(e^{-s}(1+e^{-2\beta B})C(\beta)^{-1}e^{-2\beta B} - C(\beta)^{-1}e^{-4\beta B}) \tag{IV.5.11}$$

If we let $u = e^{2\beta B}$, then our equation becomes:

$$r(s) = C(\beta)^{-1}\frac{1}{u}s\left(e^{-s}\left(1+\frac{1}{u}\right)-\frac{1}{u}\right) \tag{IV.5.12}$$

We reiterate the remark that $s \in (0, \ln(1+u))$ so that we change variables (monotonically) to $s = \ln(1+u(1-e^{-\alpha}))$, so that:

$$\tilde{r}(\alpha) = \frac{1}{C(\beta)u}\left(\ln(1+u(1-e^{-\alpha}))\left(\frac{u+1}{u(1+u(1-e^{-\alpha}))} - \frac{1}{u}\right)\right) \tag{IV.5.13}$$

$$= \frac{1}{C(\beta)u}\frac{\ln(1+u(1-e^{-\alpha})}{e^{\alpha}(1+u(1-e^{-\alpha})} \tag{IV.5.14}$$

This then gives us the bound for $|\rho|$ as:

$$|\rho| \geq C(\beta)^{-1} \max_{\alpha\in(0,\infty)} \frac{\ln(1+u(1-e^{-\alpha}))}{ue^{\alpha}(1+u(1-e^{-\alpha}))} \tag{IV.5.15}$$

If we then follow the same argument before, we get the bounds for the virial coefficients as:

$$|c_n| \leq \frac{1}{n}C(\beta)^{n-1}(\mathcal{F}(u))^{-(n-1)} \tag{IV.5.16}$$

where $\mathcal{F}(u) = \max_{a\in(0,\infty)} \frac{\ln(1+u(1-e^{-\alpha}))}{ue^{\alpha}(1+u(1-e^{-\alpha}))}$.

However, using techniques involving Canonical Ensemble calculations for free energy and its relationship as the Legendre transform of pressure, they obtain slightly better bounds on the virial coefficients, although still under the assumption that $|\rho| \leq \rho^*$, where $\rho^*$ is the radius of convergence in (IV.5.9). The improved estimates for the coefficients seem to imply that there may be a way of extending the radius of convergence, at least for temperatures for which the bound on the coefficient is an improvement. They obtain bounds on the coefficient of $\rho^{k+1}$ in the free energy as:

$$\left(\frac{1}{k+1} + (e^{\alpha^*_\beta} - 1)e^{\alpha^*_\beta k}\right) e^{2\beta B(k-1)} \frac{(k+1)^k}{k!} C(\beta)^k \tag{IV.5.17}$$

This leads to the (better) asymptotic bound for the virial coefficients as:

$$K\left(\frac{e^{2\beta B}C(\beta)}{0.24026}\right)^k \tag{IV.5.18}$$

where $K$ is a constant.

### IV.5.2 Lebowitz-Penrose

We again use the same bounds on the cluster expansion and follow the method outlined in [LePen64] and starting with (IV.5.12), we arrange our expression $r(s)$ into the form:

$$r(s) = \frac{C(\beta)^{-1}}{1+u}\frac{s}{u^2}\left((1+u)^2e^{-s} - (1+u)\right) \tag{IV.5.19}$$

We use the identity: $\frac{1+u}{u^2} = \frac{(1+u)^2}{u^2} - \frac{1+u}{u}$, to rewrite it as:

$$r(s) = \frac{C(\beta)^{-1}}{1+u}\left(s\frac{1+u}{u} - s^2\frac{(1+u)^2}{u^2}\left(\frac{1-e^{-s}}{s}\right)\right) \tag{IV.5.20}$$

We make the change of variables $v = s\frac{1+u}{u}$ and define $g(w) := \frac{1-e^{-w}}{w}$ and obtain:

$$\hat{r}(s) = \frac{C(\beta)^{-1}}{1+u}(v - v^2 g\left(\frac{uv}{1+u}\right)) \tag{IV.5.21}$$

We note that $g'(w) = \frac{we^{-w}-(1-e^{-w})}{w^2} = \frac{(w+1)e^{-w}-1}{w^2}$

If we use the inequality $e^w \geq 1 + w$, then $1 \geq (1 + w)e^{-w}$ and so $0 \geq (1+w)e^{-w}-1$ and hence $g$ is decreasing. We note that $u \in (1, \infty)$ and so $\frac{u}{u+1} \in (\frac{1}{2}, 1)$. Since $g$ is decreasing, we get maximum value at $g(\frac{v}{2})$, which would give a minimum value for the expression in brackets which gives a bound uniform in $u$. This is the extra approximation in the derivation of the virial coefficient bounds, which is not made in the general derivation. This is equivalent to realising that the $g$ function is dominated by its contribution at $u-1$, that is when $\beta = 0$, so at high temperatures.

We then seek to maximise: $f(v) = v - v^2 g(\frac{1}{2}v)$. Upon differentiation we obtain:

$$\begin{aligned} f'(v) &= 1 - 2vg(\frac{1}{2}v) - \frac{1}{2}v^2 g'(\frac{1}{2}v) \\ &= 1 - 2v\left(\frac{1-e^{-\frac{1}{2}v}}{\frac{1}{2}v}\right) - \frac{1}{2}v^2\left(\frac{(\frac{1}{2}v+1)e^{-\frac{1}{2}v}-1}{\frac{1}{4}v^2}\right) \\ &= 1 - 4(1-e^{-\frac{1}{2}v}) - 2(\frac{1}{2}v+1)e^{-\frac{1}{2}v} + 2 \\ &= -1 + 2e^{-\frac{1}{2}v} - ve^{-\frac{1}{2}v} \end{aligned} \tag{IV.5.22}$$

When we set this to zero we obtain:

$$1 = (2-v)e^{-\frac{1}{2}v} \tag{IV.5.23}$$

changing parameters to $\delta = 1 - \frac{1}{2}v$, we have:

$$\frac{e}{2} = \delta e^{\delta} \tag{IV.5.24}$$

and so using the Lambert W-function again we get:

$$\delta = W\left(\frac{e}{2}\right) \tag{IV.5.25}$$

In our original expression:

$$\begin{aligned}\tilde{f}(\delta) &= 2-2\delta-(2-2\delta)^2 g(1-\delta)\\ &= 2(1-\delta)-4(1-\delta)^2\left(\frac{1-e^{\delta-1}}{1-\delta}\right)\\ &= (1-\delta)(2-4(1-e^{\delta}e^{-1}))\\ &= 2(-1+\delta+\frac{1}{\delta}-1)\\ &= 2\frac{(W(\frac{e}{2})-1)^2}{W(\frac{e}{2})}\end{aligned} \tag{IV.5.26}$$

This therefore gives us the bound

$$|\rho| \geq \frac{2C(\beta)^{-1}}{1+u}\frac{W(\frac{e}{2})-1)^2}{W(\frac{e}{2})} \tag{IV.5.27}$$

Leading to the coefficient bound of:

$$|c_n| \leq \frac{1}{n}C(\beta)^{n-1}(1+e^{2\beta B})^{n-1}\left(\frac{W(\frac{e}{2})}{2(W(\frac{e}{2})-1)^2}\right)^{n-1} \tag{IV.5.28}$$

and the lower bound on the radius of convergence as:

$$\mathcal{R} \geq C(\beta)^{-1}\frac{1}{1+e^{2\beta B}}2\frac{(W(\frac{e}{2})-1)^2}{W(\frac{e}{2})} \tag{IV.5.29}$$

### IV.5.3 Comparison of the Bounds

If we define:

$$r_1 := e^{-4\beta B}\frac{(W\left(\frac{e}{1+e^{2\beta B}}\right)-1)^2}{W\left(\frac{e}{1+e^{2\beta B}}\right)} \tag{IV.5.30}$$

$$r_2 := \frac{1}{1+e^{2\beta B}}2\frac{(W(\frac{e}{2})-1)^2}{W(\frac{e}{2})} \tag{IV.5.31}$$

In order to concentrate on the factor over which the two bounds (IV.5.29) and (IV.5.9) differ, we need only concentrate on $r_1$ and $r_2$ above. In Figure IV.1, we see that the optimised bound shows a slight improvement over the Lebowitz-Penrose bound. Furthermore, considering the quotient $\frac{r_1}{r_2}$ in Figure IV.2, we see that the optimised bound is 1.25 times better for the low temperature limit, whereas at high temperatures the two bounds are approximately the same. This is explained by emphasising that the approximation in the $g$ function is its exact value at $\beta = 0$ or

high temperature.

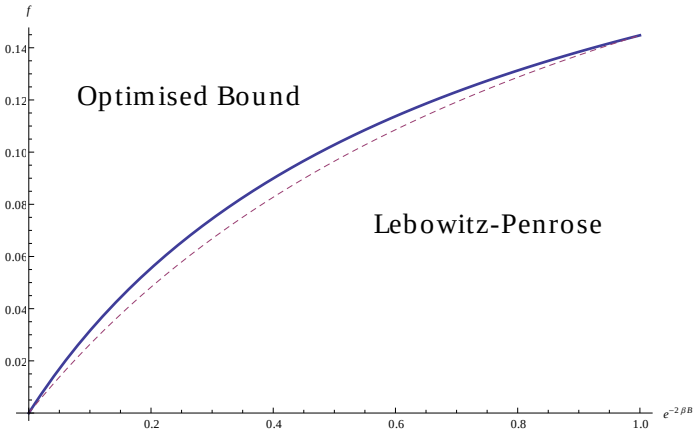



Figure IV.1: Comparison of the Lebowitz Penrose Bound ($r_1$) (IV.5.30) with My Optimised Bound ($r_2$) (IV.5.31)

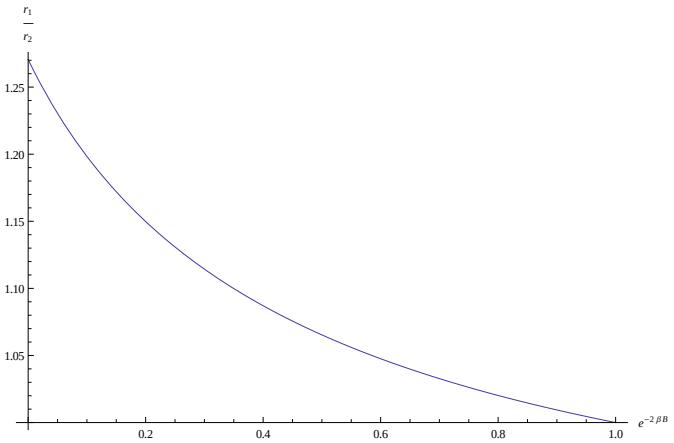



Figure IV.2: Quotient of the Optimised Bound ($r_1$) (IV.5.30) and Lebowitz Penrose Bound ($r_2$) (IV.5.31)

## IV.6 Further Bounds obtainable from the General Derivation

There are alternative cluster coefficient bounds relating to tree-graph identities found in [BrFe78, Bry84, Pro07, PoUe09]. The bounds are given by:

$$|nb_n| \leq \frac{n^{n-1}}{n!} R(\beta)^{n-1} e^{n\beta B} \tag{IV.6.1}$$

where

$$R(\beta) = \left( |\mathbb{B}| r^D + \beta \int\limits_{|y|>r} |\Phi(y)| \mathrm{d}^D y \right) \tag{IV.6.2}$$

where $|\mathbb{B}|$ denotes the surface area of the $D$-dimensional sphere.

In this case our parameters are: $a = R(\beta)^{-1}$; $b = R(\beta)e^{\beta B}$; $ab = e^{\beta B}$; and $\mu = \frac{e^{\beta B+1}}{1+e^{\beta B}}$.

The general bound for $|\rho|$ is (from (IV.4.13)):

$$|\rho| \leq R(\beta)^{-1} \frac{(W(\frac{e^{\beta B+1}}{1+e^{\beta B}}) - 1)^2}{W(\frac{e^{\beta B+1}}{1+e^{\beta B}})} \tag{IV.6.3}$$

Which gives the bound on the coefficients as:

$$|c_n| \leq \frac{1}{n} R(\beta)^{n-1} \left( \frac{W\left(\frac{e^{\beta B+1}}{1+e^{\beta B}}\right)}{(W\left(\frac{e^{\beta B+1}}{1+e^{\beta B}}\right) - 1)^2} \right)^{n-1} \tag{IV.6.4}$$

and the lower bound on the radius of convergence is:

$$\mathcal{R}_{\mathrm{Vir}} \geq R(\beta)^{-1} \frac{(W\left(\frac{e^{\beta B+1}}{1+e^{\beta B}}\right) - 1)^2}{W\left(\frac{e^{\beta B+1}}{1+e^{\beta B}}\right)} \tag{IV.6.5}$$

### IV.6.1 Comparison of the separate bound

This approach would be better for potentials where $\zeta R(\beta) < C(\beta)$, for some $\zeta$ representing the quotient of the coefficients $\frac{f_1}{f_2}$, where:

$$f_1 := e^{-4\beta B} \frac{\left(W\left(\frac{e}{1+e^{2\beta B}}\right) - 1\right)^2}{W\left(\frac{e}{1+e^{2\beta B}}\right)} \tag{IV.6.6}$$

$$f_2 := \frac{\left(W\left(\frac{e^{\beta B+1}}{1+e^{\beta B}}\right) - 1\right)^2}{W\left(\frac{e^{\beta B+1}}{1+e^{\beta B}}\right)} \tag{IV.6.7}$$

$f_2$ is actually somewhat smaller than $f_1$ and so (IV.6.5) would only be an improvement if the $R(\beta)$ factor compensates. For any potential with a hard core, this modified version is somewhat essential. Even without the hardcore, the use of the integral of $\beta|\Phi(y)|$ rather than $|e^{-\beta\Phi(t)} - 1|$ is better for negative potentials. It is also interesting to note that these coefficients have the same high and low temperature

limits.

The comparison between my optimised bound of the Lebowitz-Penrose type ($f_1$) and the bound achieved from these different cluster coefficient bounds based on tree graph identities ($f_2$) is shown in Figure IV.3. This gives us that as soon as $1.6R(\beta) < C(\beta)$ this other version of bounds is better. This is due to the fact that

$$\frac{C(\beta)}{R(\beta)} > 1.6 > \frac{f_1}{f_2} \tag{IV.6.8}$$

and so

$$f_2 R(\beta)^{-1} > f_1 C(\beta)^{-1} \tag{IV.6.9}$$

We also note that $R(\beta)$ should be a better bound, for large $\beta$ or small temperature, since it is linear in $\beta$, whereas $C(\beta)$ is exponential in $\beta$ and that this $\zeta(\beta)$ could be understood better to explore the comparison of these two bounds. It also only depends on the potential $\Phi$ through the stability parameter $B$.

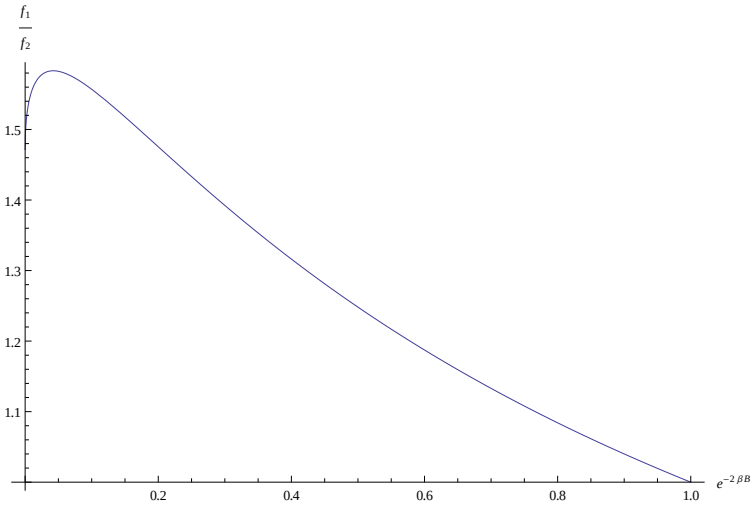



Figure IV.3: Quotient $\frac{f_1}{f_2}$ (IV.6.8) of the My Improved Penrose-Lebowitz Bound ($f_1$) (IV.6.6) over the New Bound ($f_2$) (IV.6.7)

## IV.7 Current Problems and Issues

The first main problem with the virial expansion is that the integrals are hard to compute analytically. Very few computations have been made analytically and virial coefficients up to tenth order have been computed numerically for hard disc models in [ClMc06, ClMc05, Lyb05, RH64a, RH67]. The main thrust behind these calculations is through the Ree-Hoover expansion, which is described in Section

VIII.3. This simplifies the terms needed to be calculated in the virial expansion and in the hard disk case, the number of graphs needing to be considered is reduced significantly. Furthermore, estimation of the virial coefficients is somewhat difficult. We realise that the number of two-connected graphs on $n$-points is greater than $2^{\binom{n}{2}-n}$, since we can start with the cycle on $n$ points and add (or not) any other edge to this cycle and it will be two-connected. We know that power series of the form:

$$\sum_{n=2}^{\infty} \frac{2^{n^2}}{n!} \rho^n \tag{IV.7.1}$$

have zero radius of convergence. We must control the cancellations in the weights we use for the two-connected graphs and find a way to organise the calculation of the coefficients to best take advantage of this fact. This idea is used in Chapter VIII for two-connected graphs and also in Chapter VII for connected graphs. The Penrose construction gives us an idea on how these cancellations can work in the case of connected graphs to give us tree-graph inequalities and an effective way of understanding appropriate cancellations. Furthermore, understanding the link between connected and two-connected graphs more precisely, may be able to give us a tree expression for the virial coefficients. Trees have the nice property that the number of trees on $n$ points is $n^{n-2}$, which when divided by the $n!$ on the denominator gives a non-zero radius of convergence.

Another line of approach is that of Ree and Hoover, which was taken up by McCoy and Clisby, where we try to find a way of reducing the number of graphs that need to be considered. The involutions mentioned in the previous chapter also indicate a method of how we can explore cancellations of different weights and see how terms cancel in the expansion to either ease computations or make it analytically tractable.

In the paper by Jansen [Jan12], the virial series is understood at low temperatures and in the absence of a polyatomic transition is shown to be convergent with asymptotic $\mathcal{R}_{\text{Vir}} \sim e^{-\beta B}$, which is better than the results we have which give an extra factor 2 in the exponential. A further line of investigation is whether the series expansion need no longer be valid, but the pressure or free energy function could still be analytically continuable. A challenge would be to understand the nature of this function and how the analytically continue the virial expansion to the positive density region in this case.

The current status is that we have bounds in specific cases and some general bounds. For some models we have exact results.

# Chapter V

# An Introduction to Combinatorial Species of Structure

This chapter is an introduction to the theory of species of structure, providing the key background from category theory and generating functions and their transformations. The discussion is motivated by examples from the previous chapters.

The first paper to really put the ideas of combinatorial species of structure on a firm footing in category theory, and to emphasise the connections between various forms of generating functions and combinatorial structures, particularly in the context of various operations, is the paper by Joyal [Joy81]. The ideas are certainly already somewhat present in the work by Bender and Goldman [BeGo71], as well as Bourbaki [Bou68]. Also the notion of Polya's Hauptsatz is present in the notion of cycle-index series. Joyal's paper exposes some of the powerful aspects of the theory in determining the Cayley formula for the number of trees and a version of the Lagrange inversion formula. It is in this paper that we first truly get a rigorous notion of what a labelled structure is: a functor from the category of finite sets with bijections to itself. Furthermore, it gives a good qualification of an unlabelled structure from the labelled structures via an equivalence relation. This is a notable foundational advance in the subject which overhauled the lack of precision concerning labelled structures.

Much work has been done on the subject, including natural extensions to virtual species and related combinatorial and algebraic operations and ideas. Two notable works are the book by Bergeron, Labelle and Leroux [BLL98], which contains many valuable references on the development of the subject and the book by

Flajolet and Sedgewick [FlSe09]. The important realisation of the inherent links between combinatorial species of structure and the cluster expansion and more generally in the context of statistical mechanics, via the use use of generating functions, appears to have been solidified in the late twentieth and early twenty-first centuries. The papers of Leroux [Ler04] and Faris [Far10] give a very clear idea of this link between the two subjects. Indeed Faris describes this theory of combinatorial species of structure as providing a metaphorical Rosetta Stone, from which we can translate between the combinatorial language into both versions of the cluster expansion, either in the particle picture or in the polymer model, via the use of generating functions. It is conveyed that the rewards of this approach don't just happen between combinatorics and the two versions separately, but rather it helps to unify the two different versions of the cluster expansion.

Of course the notion of combinatorics and in particular graph theory within the understanding of cluster expansions and in more general perturbation expansions was already realised by Mayer [MMay40]. The dimension that the formulation of combinatorial species of structure adds to statistical mechanics, comes partly from the general results such as the dissymmetry theorem and Lagrange inversion in simplifying and focusing the understanding of the relationships between cluster and virial expansions.

The language of combinatorial species of structure offers an essential framework in which to understand the main features of the cluster and virial expansions. It removes a lot of the focus on particular types of functions as were described in Chapter II and instead ensures that the important relations are clear and without any of the additional analytic descriptions, used in Chapter II. It is a simplification through generalisation of the key ideas behind the cluster expansion.

It is anticipated that, in the advent of this more unified approach, the combinatorial relationships can aid our interpretation and understanding of the various coefficients. This should be done in such a way that we can obtain bounds on the important expansions represented by particular species of structure. Indeed, the work of Faris [Far10] has made a great initiation of this idea, by providing how fixed point equations can be developed separately through the framework of species of structure. These give rise to the common conditions that are used and developed for obtaining bounds on cluster expansions.

Indeed, one can continue with this explicit interpretation of what the main theorems in cluster expansions represent and mean in terms of combinatorial species of structure. Most of the ideas are pretty strong for cluster expansions and connected graphs. It is anticipated that these relationships can provide inspiration for

relationships for virial coefficients and two-connected graphs.

## V.1 Definition of Combinatorial Species

### V.1.1 Some Category Theory

The use of category theory to give a clear definition for algebraic structures is presented to introduce the combinatorial species of structure. Introductory books used for this section on category theory are [Sim11, Mac98]. The definitions for combinatorial species of structure can be found in [Joy81, BLL98, FlSe09]. This section aims to provide motivation and diagrams to emphasise the key points of these definitions and their application.

The main thrust of the work by Joyal [Joy81], is to give a precise definition of a labelled structure, through a definition as a functor between the category of finite sets with bijections and itself.

**Definition.** *A Category $C$ consists of:*

1. *A class $Ob(C)$ whose elements are called objects*
2. *A class $hom(C)$ whose elements are called morphisms. Each morphism has a source object $a$ and target object $b$. The set of morphisms from $a$ to $b$ is denoted $hom(a, b)$*
3. *A binary operation $\circ$, called the composition of morphisms. For any three objects $a$, $b$ and $c$, we have:*

$$\circ : hom(a, b) \times hom(b, c) \to hom(a, c) \tag{V.1.1}$$

*We write the composition $f \circ g$ or $fg$. This has to satisfy the two properties:*

   (a) *Associativity: If $f : a \to b$, $g : b \to c$ and $h : c \to d$, then $h \circ (g \circ f) = (h \circ g) \circ f$*
   (b) *Identity: For every object $x$, there exists a morphism $I_x : x \to x$ called the identity morphism, such that for each $f : a \to b$, $f \circ I_a = I_b \circ f = f$*

Categories also contain the natural idea of a functor: a structure preserving map between categories.

**Definition.** *A (covariant) functor from a category $C$ to a category $D$ written $F : C \to D$, consists of:*

1. *For each object $x$ in $C$ an object $F(x)$ in $D$*

2. *For each morphism $f : x \to y$ in $C$ a morphism $F(f) : F(x) \to F(y)$ in $D$*

*such that the following two properties hold:*

1. *For every object $x$ in $C$, $F(Id_x) = Id_{F(x)}$*

2. *For all morphisms $f : x \to y$ and $g : y \to z$, $F(g \circ f) = F(g) \circ F(f)$*

### V.1.2 The Definitions

Progress on enumerating particular structures on finite sets has been approached by many methods. Some of these methods are:

1. through relations between already enumerated structures

2. through implicit functional equations

3. through recursive relationships - building a structure from other structures on smaller sets

4. through the use of the inclusion-exclusion principle and Möbius formulæ

The idea to write an exponential generating function, with coefficients being the number of such structures on a set of size $n$, was developed to allow for the techniques of asymptotic analysis to be used to understand asymptotic enumeration. This is given a thorough exposition in the book by Flajolet and Sedgewick [FlSe09].

The key power of combinatorial species of structure is to relate combinatorial operations to operations on (exponential) power series. Furthermore, generating functions are a natural concept in probability, where the coefficients represent moments of a random variable. The links between probability and combinatorics are strengthened by this relationship.

A combinatorial species of structure is a functor from the category of finite sets with bijections to itself, as described below:

**Definition.** *A species of structures is a rule $F$, which*

*i) Produces for each finite set $U$ a finite set $F[U]$*

*ii) Produces for each bijection $\sigma : U \to V$ a bijection $F[\sigma] : F[U] \to F[V]$*

*The functions $F[\sigma]$ are required to satisfy the following functorial properties:*

*i) For all bijections $\sigma : U \to V$ and $\tau : V \to W$, We have: $F[\tau \circ \sigma] = F[\tau] \circ F[\sigma]$*

*ii) For the identity map $Id_U : U \to U$, we have $F[Id_U] = Id_{F[U]}$*

*An element $s \in F[U]$ is called an $F$-structure on $U$.*

**Example** (Examples of Species of Structure)**.** *Five main non-graphical examples of species of structure used in this thesis are:*

1. *The SET species $\mathcal{E}$, where for every finite set $U$ we have the set of species of structures is $\{U\}$.*
2. *The PER species $\mathcal{S}$, where for each finite set we have the structures $S_{|U|}$- the set of all permutations on $U$.*
3. *Species of structure $X$. For any finite set $U$, we get $\{U\}$ if $|U| = 1$ and $\emptyset$ otherwise.*
4. *The Species $\mathcal{E}_N$ - the indicator species of sets of size $N$. We have $\mathcal{E}_N[U] = \{U\}$ if $|U| = N$ and $\emptyset$ otherwise.*
5. *The power set or subset species $\mathcal{P}$, gives for $U$ all subsets of $U$.*

The most useful generating function for statistical mechanics is the exponential generating function:

**Definition.** *The exponential generating series of a species of structure $F$ is the formal power series:*

$$F(x) = \sum_{n=0}^{\infty} f_n \frac{x^n}{n!} \tag{V.1.2}$$

*where $f_n = |F[n]|$, the number of (labelled) $F$-structures on a set of $n$ points.*

**Example** (Exponential Generating Series for our Examples)**.** *For our examples above, the exponential generating series are:*

1. *For SET $\mathcal{E}$, we have that $|\mathcal{E}[n]| = 1\ \forall n$, so $E(x) = \sum_{n=0}^{\infty} \frac{x^n}{n!} = \exp(x)$*
2. *For PER $\mathcal{S}$, we have $|\mathcal{S}[n]| = n!\ \forall n$, so $S(x) = \sum_{n=0}^{\infty} x^n = \frac{1}{1-x}$*
3. *For $X$, we have $|X[1]| = 1$ and $|X[n]| = 0$ for $n \neq 1$, so $X(x) = x$*
4. *For $\mathcal{E}_N$, we have $|\mathcal{E}_N[N]| = 1$ and $|\mathcal{E}_N[n]| = 0$ for $n \neq N$, so $E_N(x) = \frac{x^N}{N!}$*
5. *For $\mathcal{P}$, we have $|\mathcal{P}[n]| = 2^n$ and so $P(x) = \sum_{n=0}^{\infty} \frac{2^n}{n!} x^n = \exp(2x)$*

There are two other important related generating functions for species of structure.

If we have unlabelled species of structure, then we don't have to divide by the $n!$ for permuting labels and we use the ordinary generating function. The isomorphism type of a species of structure is an equivalence relation on $F[n]$ with

$$s \sim t \text{ if } \exists \pi : [n] \to [n] \text{ such that } F[\pi](s) = t \tag{V.1.3}$$

**Definition** (The Isomorphism Type Generating Series)**.** *The isomorphism type generating series of a species of structure $F$ is the formal power series:*

$$\tilde{F}(x) = \sum_{n=1}^{\infty} \tilde{f}_n x^n \tag{V.1.4}$$

*where $\tilde{f}_n$ is the number of isomorphism types (equivalence classes) of $F$-structures on $[n]$.*

A further generating series which contains a lot more information about our species of structure is the cycle index series $Z_F$. For a general permutation $\sigma$, its cycle type is the sequence $(\sigma_1, \sigma_2, \cdots)$, where $\sigma_k$ is the number of cycles of length $k$ in the decomposition of $\sigma$ into disjoint cycles. We define the quantities:

$$\text{Fix } \sigma := \{u \in U | \sigma(u) = u\} \quad \text{and} \quad \text{fix } \sigma = |\text{Fix } \sigma| \tag{V.1.5}$$

**Definition** (Cycle Index Series)**.** *The cycle index series of a species of structure $F$ is the formal power series in infinitely many variables:*

$$Z_F(x_1, x_2, \cdots) = \sum_{n \geq 0} \frac{1}{n!} \left( \sum_{\sigma \in S_n} \mathit{fix}\ F[\sigma] x_1^{\sigma_1} x_2^{\sigma_2} \cdots \right) \tag{V.1.6}$$

This is related to Polya's Hauptsatz, which is applied to Husimi trees and enumerating connected graphs in terms of two-connected graphs in the work of Ford and Uhlenbeck [FoUh56a, FNU56, FoUh56b, FoUh57].

**Definition.** *Let $F$ and $G$ be two species of structure. An equipotence $\alpha : F \to G$ is a family of bijections $\alpha_U : F[U] \to G[U]$, for each finite set $U$.*

*If there exists an equipotence between two species $F$ and $G$, then the two species are called equipotent, which we write as $F \equiv G$.*

**Remark 28.** *We note that $F(x) = G(x)$ is equivalent to $F \equiv G$*

The idea of combinatorial equality is somewhat more restrictive:

**Definition.** *An isomorphism of $F$ to $G$ is a family of bijections $\alpha_U : F[U] \to G[U]$, which obeys the naturality condition:*

*For any $F$-structure $s \in F[U]$ and $\sigma : U \to V$ bijection, one must have*

$$\sigma \circ \alpha_U(s) = \alpha_V(\sigma \circ s) \tag{V.1.7}$$

*That is the following diagram commutes:*

$$\begin{array}{ccc} F[U] & \xrightarrow{\alpha_U} & G[U] \\ {\scriptstyle F[\sigma]}\downarrow & & \downarrow{\scriptstyle G[\sigma]} \\ F[V] & \xrightarrow{\alpha_V} & G[V] \end{array}$$

*The two species are then said to be isomorphic and one writes $F \simeq G$*

Isomorphic species have the same cycle index series and isomorphism type generating series. Equipotent species don't necessarily share the same series.

## V.2 Operations on Species of Structure

### V.2.1 Addition

As a simple motivating example, we have the splitting of graphs into connected and disconnected graphs, so we would want to be able to have the natural expression:

$$\mathcal{G} = \mathcal{C} + \mathcal{G}^d \tag{V.2.1}$$

The idea of 'summing' two species of structure is quite natural. We may, for example have two different types of graph such as cycles and complete graphs and then we wish to understand the species comprised of both of these. In some sense this idea of summing species of structure is like aggregating different species into a whole species. We wish to understand a whole group of structures, by summing out the individual details of the smaller structures. Let $F$ and $G$ be two species of structure:

**Definition.** *The sum of $F$ and $G$, denoted $F + G$,is defined as follows:*

*An $F + G$ -structure on a finite set $U$ is either an $F$-structure on $U$ or (exclusive) a $G$ structure on $U$.*

*In other words, for any finite set $U$, we have that $(F+G)[U] = F[U]+G[U]$, where the $+$ indicates the disjoint union operation.*

*The transport along the bijection $\sigma : U \to V$ is carried out by defining for any $(F+G)$-structure $s$ on $U$:*

$$(F+G)[\sigma](s) = \begin{cases} F[\sigma](s) & \text{if } s \in F[U] \\ G[\sigma](s) & \text{if } s \in G[U] \end{cases} \tag{V.2.2}$$

**Remark 29.** *Note that under this definition $(F+G)(x) = F(x) + G(x)$, since $|(F+G)[n]| = |F[n]| + |G[n]|$. A natural extension to the sum of two species of structure is to be able to sum an arbitrary collection of species. The case when we can do this is given precisely by what is defined by the word 'summable'.*

**Definition** (Summable Collection of Species of Structure)**.** *A family $(F_i)_{i\in I}$ of species of structure is said to be summable, if for any finite set $U$, $F_i[U] = \emptyset$, except for a finite number of indices $i \in I$. The sum of a summable family $(F_i)_{i\in I}$ is the species $\sum_{i\in I} F_i$ defined by:*

$$\left(\sum_{i\in I} F_i\right)[U] = \sum_{i\in I} F_i[U] \times \{i\} \tag{V.2.3}$$

$$\left(\sum_{i\in I} F_i\right)[\sigma](s,j) = (F_j[\sigma](s), j) \tag{V.2.4}$$

*where the second line gives the transport under the bijection $\sigma : U \to V$ and*

$$(s,j) \in \left(\sum_{i\in I} F_i\right)[U] \tag{V.2.5}$$

Figure V.1 emphasises that if our $F$ and $G$-structures overlap (in this case with the 4-cycle), then the $F+G$-structure contains this structure twice. The vertices in this example are coloured differently to emphasise the fact that we treat the summed structures as different if they come from different initial species.

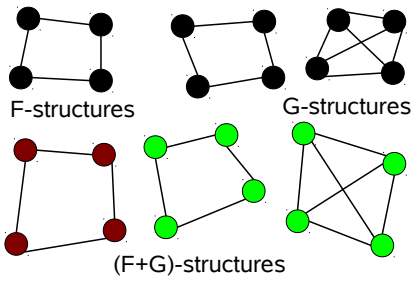



Figure V.1: The sum of two species of structure

### V.2.2 Multiplication

Multiplication is a natural operation of power series. The 'star' product, which is used in the Mathematical Physics literature to indicate products of graphs, but where, in the integral expression corresponding to the graph, we need to add an appropriate 'symmetry factor'. For example, in [Ste65], Stell introduces the factors in order to get the 'combinatorics' of multiplication correct. This comes from the problem of labelled structures, where the ability to relabel a structure, but not change it, gives particular issues. For example in Figure V.2, we see that we can swap the labels 2 and 3 without changing the actual labelled tree. This means the graph would obtain a symmetry factor of $\frac{1}{2}$, whereas individual edges have symmetry factor 1 and so they don't multiply correctly.

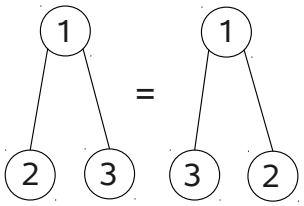



Figure V.2: Two Equal Simple Trees

The notion of combinatorial species of structure and its multiplication operations bypasses such issues as the whole set up is there to automatically deal with the symmetry factors. The natural way to deal with such an issue is to set labels for each structure you start with, which is equivalent to partitioning the label set $U$ into two sets $U_1 \cup U_2 = U$ $U_1 \cap U_2 = \emptyset$ and construct the separate species on the subsets.

Let $F$ and $G$ be two species of structure:

**Definition.** *The product species $F \star G$ called the product of $F$ and $G$, is defined as follows: an $(F \star G)$-structure on $U$ is an ordered pair $s = (f, g)$, where:*

*i) $f$ is an $F$-structure on a subset $U_1 \subset U$*

*ii) $g$ is a $G$-structure on a subset $U_2 \subset U$*

*iii) $(U_1, U_2)$ is a decomposition of $U$ i.e. $U = U_1 \cup U_2$ and $U_1 \cap U_2 = \emptyset$.*

*In other words, for any finite set $U$:*

$$(F \star G)[U] = \sum_{(U_1,U_2)} F[U_1] \times G[U_2] \tag{V.2.6}$$

*The sum being taken over all pairs $(U_1, U_2)$ forming a decomposition of $U$.*

*The transport along a bijection $\sigma : U \to V$ is carried out by setting, for each $(F \star G)$-structure $s = (f, g)$ on $U$, $(F \star G)[\sigma](s) = (F[\sigma_1](f), G[\sigma_2](g))$, where $\sigma_i = \sigma|_{U_i}$, the restriction of $\sigma$ on $U_i$, $i = 1, 2$.*

We note that:

$$(F \star G)(x) = \sum_{n=1}^{\infty} \sum_{m=1}^{n} \frac{x^n}{n!} \binom{n}{m} f_m g_{n-m} = \sum_{n=1}^{\infty} \sum_{m=1}^{n} \frac{x^n}{m!(n-m)!} f_m g_{n-m} \tag{V.2.7}$$

We also have:

$$\begin{aligned} F(x)G(x) &= \sum_{m=1}^{\infty} \frac{x^m}{m!} f_m \sum_{k=1}^{\infty} \frac{x^k}{k!} g_k \\ &= \sum_{n=1}^{\infty} \sum_{m=1}^{n} \frac{x^n}{m!(n-m)!} f_m g_{n-m} \\ &= (F \star G)(x) \end{aligned} \tag{V.2.8}$$

In Figure V.3, we see how to multiply cycles with complete graphs. It shows one of the structures we may obtain on a vertex set of size 10.

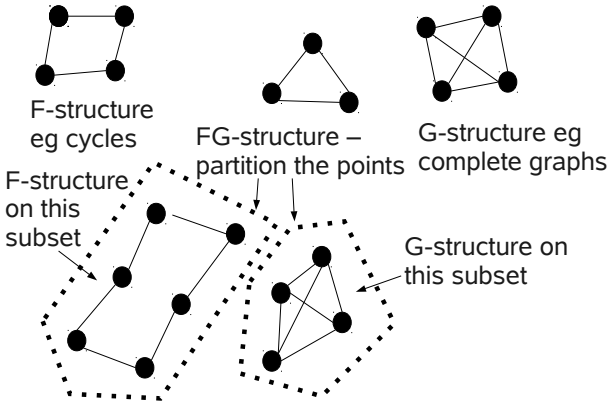



Figure V.3: The Product of Two Species of Structure

The idea of multiplication can be extended many times to understand products of the form $\mathcal{K}^k$, where $\mathcal{K}$ is a generic species of structure. This would involve making a partition of the original label set in $k$ subsets (a $k$-partition) and having separate $\mathcal{K}$-structures on each, where it is important to note that we have an 'order' to the partition, due to our definition of product. If we wish to neglect what order our $\mathcal{K}$-structures are in then we need to divide by $k!$. This is revisited in Section VI.2 on Lagrange inversion. The key idea is that the partitions obtained from multiplication are necessarily ordered, since we have to include the case where the structures are different. However, if we want to understand separate identical structures, we need to divide by the factorial.

In order to generalise the splitting into two sets, we take appropriate powers of the species of structure, by repeating the multiplication operation. If all possibilities of partitioning a set is required, then one needs to sum the series of products over factorials. A technical point is we need to make sure that $\mathcal{K}[\emptyset] = \emptyset$ in order to do this sum or else we can take the $\emptyset$ in our partition, which we do not want (for convergence). If we want any partition, where each set in the partition has an $\mathcal{A}$-structure, then we need the species:

$$\sum_{k=0}^{\infty} \frac{\mathcal{K}^k}{k!} = \exp(\mathcal{K}) \tag{V.2.9}$$

We know that the exponential generating function of SET $\mathcal{E}$ species is $\exp(x)$ and so this looks like a composition. This leads us onto defining a notion of composition/substitution.

### V.2.3 Substitution

Composition of generating functions is a very natural operation and the combinatorial analogue is to see the external species as providing a global structure between smaller components, having the structure of the internal species.

A simple motivating example is the ability to understand permutations from disjoint cycles. The universal structure is to leave things as they are (the are no connections between disjoint cycles - this corresponds to what is known as the $SET$ species) and the local internal structures are cycles. If we follow this idea, we get:

$$PER = SET \circ CYC \tag{V.2.10}$$

More formally this conveys the idea of partitioning our set of labels. The structure between partition classes is the 'global' structure and the structure inside each set in the partition is the 'internal' structure. If we wish to only obtain the partition and have no 'connections' between different partition classes, then we use the $SET$ species. One further example of this idea is in the notion of graphs and connected graphs. We can see, similar to splitting a permutation into disjoint cycles, that a graph can be understood as its collection of connected components. We thus have the relationship:

$$\mathcal{G} = SET \circ \mathcal{C} \tag{V.2.11}$$

We now make these ideas more precise:

Let $F$ and $G$ be two species of structure, such that $G[\emptyset] = \emptyset$:

**Definition.** *The species $(F \circ G)$, also denoted $F(G)$ is called the (partitional) composite of $G$ in $F$. It is defined as follows: an $(F \circ G)$-structure on a finite set $U$ is a triplet $(\phi, \pi, \gamma)$, where:*

*i) $\pi$ is a partition of $U$*

*ii) $\phi$ is an $F$-structure on the set of classes of $\pi$*

*iii) $\gamma = (\gamma_P)_{P \in \pi}$, where for each class $P$ of $\pi$ we have $\gamma_P$ is a $G$-structure on $P$.*

*In other words for any finite set $U$, one has:*

$$(F \circ G)[U] = \sum_{\pi partition\ of\ U} F(\pi) \times \prod_{P \in \pi} G[P] \tag{V.2.12}$$

*The disjoint sum being taken over the set of partitions $\pi$ of $U$.*

*The transport along a bijection $\sigma : U \to V$ is carried out by setting, for any $(F \circ G)$-structure $s = (\pi, \phi, (\gamma_P)_{P \in \pi})$ on $U$, $(F \circ G)[\sigma](s) = (\bar{\pi}, \bar{\psi}, (\bar{\gamma}_{\bar{P}})_{\bar{P} \in \bar{\pi}})$, where:*

*i)* $\bar{\pi}$ *is the partition of* $V$ *obtained by transport along* $\sigma$ *of the partition* $\pi$

*ii) for each* $\bar{P} = \sigma(P) \in \bar{\pi}$*, the structure* $\bar{\gamma}_{\bar{P}}$ *is obtained from the structure* $\gamma_P$ *by* $G$ *-transport along* $\sigma|_P$

*iii) The structure* $\bar{\phi}$ *is obtained from* $F$*-transport along the bijection* $\bar{\sigma}$ *induced on* $\pi$ *by* $\sigma$*.*

*We can summarise this by saying an* $(F \circ G)$*-structure is an* $F$*-assembly of* $G$*-structures.*

We note that:

$$(F \circ G)(x) = \sum_{n=0}^{\infty} \frac{x^n}{n!} \sum_{m=0}^{n} \frac{f_m}{m!} \sum_{\substack{(\pi_1,\cdots,\pi_m) \\ \pi_i \subset [n] \text{ and } \cup_{i=1}^{m} \pi_i = [n] \\ \text{and } \pi_i \cap \pi_j = \emptyset \text{ for } i \neq j}} \prod_{i=1}^{n} g_{|\pi_i|} \tag{V.2.13}$$

$$= \sum_{n=0}^{\infty} \frac{x^n}{n!} \sum_{m=0}^{n} \frac{f_m}{m!} \sum_{\substack{(k_1,\cdots,k_m) \\ \sum_{i=1}^{m} k_i = n}} n! \prod_{i=1}^{m} \frac{g_{k_i}}{k_i!}$$

$$= \sum_{m=0}^{\infty} \frac{f_m}{m!} \sum_{n=m}^{\infty} \prod_{i=1}^{m} \left( \sum_{k_i=1}^{\infty} \frac{g_{k_i} x^{k_i}}{k_i!} \right) \mathbb{1}_{\sum_{i=1}^{m} k_i = n}$$

$$= \sum_{m=0}^{\infty} \frac{f_m}{m!} \left(G(x)\right)^m$$

$$= F(G(x)) \tag{V.2.14}$$

The first $\frac{1}{m!}$ comes from taking the partition as a sequence rather than a set for conveniences of computation. Hence we have the corresponding identity for generating functions. We note that this is just a simple generalisation of the derivation of the pressure coefficients representing connected graphs as in (I.2.16) and that now it has been shown in generality for labelled structures with exponential generating functions, the result can easily be applied to many other related structure to obtain exponential generating functions and their relationships.

In Figure V.4, we see an illustration of and $F(G)$-structure. The $F$-structure is conveyed between the partition sets and the $G$ structures are within the sets.

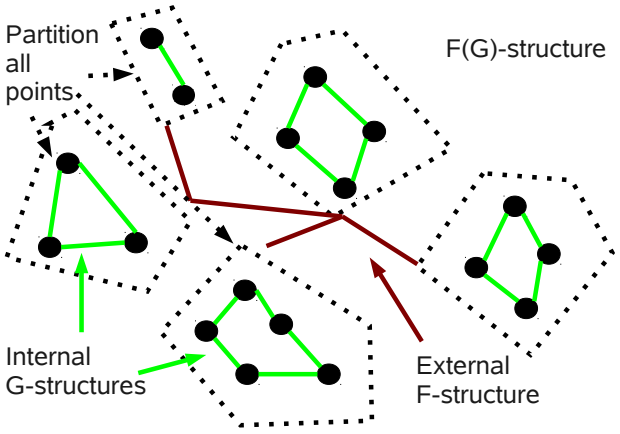



Figure V.4: The composition of two species of structure

### V.2.4 The Derivative Structure

We now have a rich algebraic structure to our combinatorial species reflecting the structure of cluster expansions. There are two important relations which are left on labelled structures. In combinatorics, we often have recursive relationships on the number of structures and it is therefore useful to have a method of 'shifting' the sequence $(k_n)$, where $k_n$ denotes the number of structures on $n$ elements of species $\mathcal{K}$. This shifting is related to formal differentiation of power series. Furthermore, we may also like to have 'rooted' structures , where a label is distinguished from the others. The main motivation for including a root comes from the structure of trees. In Computer Science, tree structures often have a root or starting point at the top of the tree or in genealogy a 'single' ancestry from which everyone can trace back their history. This notion of a root is very natural and is related to the derivative of a power series.

**Definition.** *The species $F'$ is called the derivative of the species $F$ and is defined as follows:*

*An $F'$-structure on $U$ is an $F$-structure on $U^+ = U \cup \star$, where $\star = \star_U$ is an element chosen outside of $U$. In other words, for any finite set $U$ , one sets $F'[U] = F[U^+]$, where $U^+ = U + \star$. The transport along a bijection $\sigma : U \to V$ is carried out by setting, for any $F'$-structure $s$ on $U$, $F'[\sigma](s) = F[\sigma^+](s)$, where $\sigma^+ : U + \star \to V + \star$ is the canonical extension of $\sigma$ obtained by setting: $\sigma^+(u) = \sigma(u)$ for $u \in U$ and $\sigma^+(\star) = \star$.*

**Remark 30.** *We note that*

$$F'(x) = \sum_{n=0}^{\infty} \frac{x^n}{n!} f_{n+1} = \frac{d}{dx} F(x) \tag{V.2.15}$$

Figure V.5 displays a structure on the labelled set plus an extra 'ghost' vertex, which carries the 'ghost' label rather than a label from the defining set.

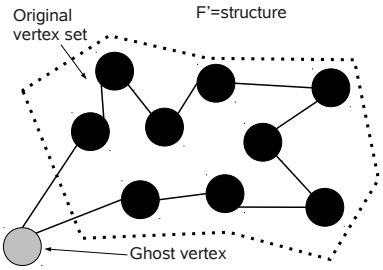



Figure V.5: The Derivative of Species of Structure

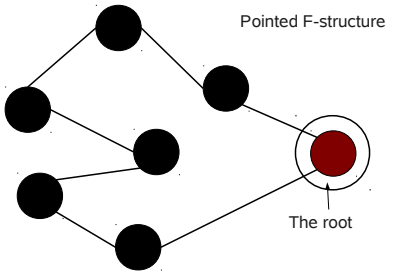



Figure V.6: A Pointed Species of Structure

### V.2.5 Rooted or Pointed Structures

**Definition.** *The pointed species $F^\bullet$ is defined as follows: An $F^\bullet$-structure on $U$ is a pair $s = (f, u)$, where:*

*i) $f$ is an $F$-structure on $U$*

*ii) $u$ is a distinguished point in $U$*

*The pair $(f, u)$ is called a pointed $F$-structure (pointed at the distinguished element $u$). In other words for any finite set $U$, $F^\bullet[U] = F[U] \times U$.*

*The transport along the bijection $\sigma : U \to V$, is carried out by setting $F^\bullet[\sigma](s) = (F[\sigma](f), \sigma(u))$, for any $F^\bullet$-structure $s = (f, u)$ on $U$.*

**Remark 31.** *We note that the pointed species of structure corresponds to the Euler derivative of the generating function.*

$$F^\bullet(x) = \sum_{n=1}^{\infty} n \frac{x^n}{n!} f_n = x \frac{d}{dx} F(x) \tag{V.2.16}$$

Figure V.6 illustrates how a point in the set is chosen to be the 'root', but otherwise we just have the original structure.

### V.2.6 Cartesian Product of Species

A natural operation on sets is to take the Cartesian product of two sets. If we do this with species then we have:

**Definition.** *The Cartesian Product of two species of structure $F$ and $G$, denoted $F \times G$, on the finite set $U$, is the collection of structures $(f, g)$, where $f$ is an $F$-structure on $U$ and $g$ is a $G$-structure on $U$.*

The action on the exponential generating functions doesn't have a nice relation, other than:

$$F \times G(x) = \sum_{n=0}^{\infty} \frac{f_n g_n}{n!} x^n \tag{V.2.17}$$

This is the Hadamard Product, but it isn't a natural operation.

### V.2.7 Functorial Composite

As a functor between categories we can understand the corresponding functorial composition of two species of structure. A notable expression in terms of the functorial composition is we can express the collection of graphs as:

$$\mathcal{G} = \mathcal{P} \square \mathcal{P}^{(2)} \tag{V.2.18}$$

This is that a graph is a subset (represented by the structure $\mathcal{P}$ - the set of all subsets or 'power set') of edges, which is the collection of two-subsets of our original finite set $U$, which is given by $\mathcal{P}^{(2)}$.

**Definition** (Functorial Composite)**.** *The functorial composite of two species of structure $F$ and $G$, is denoted $F \square G$ and defined by:*

*$F \square G[U] = F[G[U]]$ - i.e. it is all the possible $F$-structures on the collection of $G$-structures.*

*The transport is simply the composition of transports: $(F \square G)[\sigma] = F[G[\sigma]]$.*

The corresponding exponential generating function is:

$$(F \square G)(x) = \sum_{n=0}^{\infty} f_{g_n} \frac{x^n}{n!} \tag{V.2.19}$$

In the algebra of functorial composition, the neutral element is the 'pointed set' species $\mathcal{E}^{\bullet}$.

That is: $F\Box\mathcal{E}^\bullet = \mathcal{E}^\bullet\Box F = F$. This is because the species $\mathcal{E}^\bullet$ takes any finite set $U$ and gives $\mathcal{E}^\bullet[U] = U$, by identifying each pointed set by the element it is pointed at.

## V.3 Weighted Species

In many applications of the theory of species of structure, we may want different weightings for our objects that are not simply unity. For example in probabilistic contexts, we may want to represent the value of a structure as its probability of occurring. If we have the set $\mathcal{K}[n]$, we may want to put a uniform weight on each object and so need to change the $\frac{1}{n!}$ to $\frac{1}{k_n}$. We may also want to use parameters to represent particular features of the objects. In statistical mechanics, as we have seen in Chapter I, Mayer's weight for graphs is:

$$w(g) := \prod_{\{i,j\}\in E(g)} f_{i,j} \tag{V.3.1}$$

or the integral of this quantity $W(g)$.

Adding weights thus provides the ability to extend the previous notions beyond those to do with counting. Care must be taken that the weights act as we would like over the operations we consider.

**Definition.** *For a ring $\mathbb{A}$, an $\mathbb{A}$-weighted set is a pair $(A, w)$, where $A$ is a set and $w : A \to \mathbb{A}$ is a function. which associates a weight $w(\alpha) \in \mathbb{A}$ to each $\alpha \in A$*

**Definition.** *The inventory of a weighted set $(A, w)$ is defined as:*

$$|A|_w = \sum_{\alpha\in A} w(\alpha) \tag{V.3.2}$$

**Definition.** *Let $(A, w)$ and $(B, v)$ be $\mathbb{A}$-weighted sets. A morphism of $\mathbb{A}$-weighted sets $f : (A, w) \to (B, v)$ is a function $f : A \to B$, compatible with the weights, that is to say $w = v \circ f$. Moreover, if $f$ is a bijection, $f$ is called an isomorphism of weighted sets and we write $(A, w) \simeq (B, v)$.*

**Definition.** *Let $(A, w)$ and $(B, v)$ be $\mathbb{A}$-weighted sets. Define:*

*i) The sum $(A, w)+(B, v)$ as the $\mathbb{A}$-weighted set $(A+B, \mu)$, where $\mu$ is the weight function defined by:*

$$\mu(x) = \begin{cases} w(x) & \text{if } x \in A \\ v(x) & \text{if } x \in B \end{cases} \tag{V.3.3}$$

*ii) The product* $(A, w) \times (B, v)$ *as the* $\mathbb{A}$*-weighted set* $(A \times B, \rho)$*, where* $A \times B$ *denotes the Cartesian product of sets* $A$ *and* $B$ *and* $\rho$ *is the weight function defined by:* $\rho(x, y) = w(x)v(y)$*.*

**Definition** (Weighted Species)**.** *Let* $\mathbb{A}$ *be a ring of formal power series or of polynomials over a ring* $\mathbb{K} \subset \mathbb{C}$*. An* $\mathbb{A}$*-weighted species is a rule* $F$*, which:*

*i) produces for each finite set* $U$*, a finite or summable* $\mathbb{A}$*-weighted set* $(F[U], w_U)$*.*

*ii) produces, for each bijection* $\sigma : U \to V$*, a morphism* $F[\sigma] : (F[U], w_U) \to (F[V], w_V)$

*Furthermore the functions* $F[\sigma]$ *must satisfy the following functoriality properties:*

*i) If* $\sigma : U \to V$ *and* $\tau : V \to W$ *are bijections, then* $F[\tau \circ \sigma] = F[\tau] \circ F[\sigma]$

*ii) For each set* $U$*, if* $Id_U$ *denotes the identity bijection of* $U$*, then* $F[Id_U] = Id_{F[U]}$

**Definition.** *Let* $F = F_w$ *be an* $\mathbb{A}$*-weighted species of structure. The exponential generating series of* $F$ *is the exponential power series* $F_w(x)$ *with coefficients in* $\mathbb{A}$*, defined by:*

$$F_w(x) = \sum_{n \geq 0} |F_w[n]| \frac{x^n}{n!} \tag{V.3.4}$$

*where* $|F_w[n]|$ *is the inventory of the set of* $F$*-structures on* $[n]$*.*

Below is a helpful table summarising how weights are affected by particular combinatorial operations described thus far:

Table V.1: Weights for Operations on Species of Structures

| Species | Structure | Weight |
|---|---|---|
| $F_w + G_v$ | $s$ | $w(s)$ if $s \in F[U]$<br>$v(s)$ if $s \in G[U]$ |
| $F_w G_v$ | $s = (f, g)$ | $w(f)v(g)$ |
| $F_w \circ G_v$ | $s = (\pi, f, (\gamma_p)_{p \in \pi})$ | $w(f) \prod_{p \in \pi} v(\gamma_p)$ |
| $F'_w$ | $s$ | $w(s)$ |
| $F^{\bullet}_w$ | $s = (f, u)$ | $w(f)$ |

Thus to extend species isomorphisms to weighted isomorphisms, one requires the weights to behave in the way indicated in the above table.

## V.4 Graph Theory

The key examples of combinatorial species of structure relevant for cluster and virial expansions are graphical species. The key structures, we have seen are simple graphs $\mathcal{G}$, connected graphs $\mathcal{C}$, two-connected graphs $\mathcal{B}$ and trees $\mathfrak{a}$. The combinatorial relationships we can find between these structures can also be applied to the generating functions, which have an interpretation in statistical mechanics.

In Section I.2, we obtain pressure as the sum over weighted connected graphs using the combinatorial relation:

$$\mathcal{G} = \mathcal{E} \circ \mathcal{C} \tag{V.4.1}$$

We can now understand this more readily from the above theory, firstly by realising it as a combinatorial identity, since a graph is a partition of its vertex set with a connected graph on each set in its partition and then by realising the necessary requirement of the graph weights. Here it is that they factorise over connected components. This is readily understood from the integrand, where there are no functions depending on a coordinate in one component and a coordinate in another and so they must factorise into the integrals over each connected component separately.

In Section I.7, we have $D$-functions and their relationship with the connected functions, which can be succinctly written as:

$$\mathcal{B}'(\mathcal{C}^\bullet) = \mathcal{P}_+ \times \mathcal{C} \tag{V.4.2}$$

To understand this relationship, we realise that a $\mathcal{B}'(\mathcal{C}^\bullet)$-structure imposes the condition that the ghost vertex, in the derivative of the $\mathcal{B}$-structure, is not an articulation point. If we remove the ghost point from the graph on the left hand side we still have a connected graph on the non-ghost points. On the right hand side this connected graph structure is represented by $\mathcal{C}$. We may have any non-empty subset of the vertex set attached to the ghost vertex, which is indicated by the $\mathcal{P}_+$-structure, which selects a non-empty subset. It is a Cartesian product as we have both a $\mathcal{P}_+$ and a $\mathcal{C}$-structure on the *whole* set. If we wish to give both sides Mayer edge weights, where the ghost vertex is represented by label 1 and the other vertices are in the label set $[2, n]$, then we make the connection that $\mathcal{D}$-graphs are precisely the $\mathcal{B}'(\mathcal{C}^\bullet)$-graphs, with the $D$ function corresponding precisely to the Mayer-weighted graph, identifying 1 as the ghost vertex. The weight function for connected graphs is precisely the Ursell function $U$, which is necessarily on vertex set $[2, n]$, as identified above. The Mayer edge weights for the non empty connection

from vertex 1 to the set $[2, n]$ is of the form:

$$\prod_{j=2}^{n}(1 + f_{1,j}) - 1 \tag{V.4.3}$$

The first term indicates all of the possibilities of a subset of edges between 1 and $[2, n]$ and the second term removes the possibility of no edges. We thus obtain the identity:

$$D_n(x)_n = \left(\prod_{j=2}^{n}(1 + f_{1,j}) - 1\right) U_{n-1}(x)'_n \tag{V.4.4}$$

We can understand a graph with a ghost vertex - the species $\mathcal{G}'$- as consisting of a partition of the underlying set into that which is connected to the ghost vertex and that which is not. The set of points connected to the ghost vertex has, by definition, a $\mathcal{C}'$-structure and the rest of the vertices have a generic $\mathcal{G}$-structure. This leaves us with the combinatorial equation:

$$\mathcal{G}' = (\mathcal{C}' \star \mathcal{G}) \tag{V.4.5}$$

We understand simple graphs in terms of their connected components and see that the relationship is $\mathcal{G} = \mathcal{E} \circ \mathcal{C}$, since we get just the partitions with this composition.

Mayer's Second Theorem is put into this combinatorial framework in Section VI.3, where the dissymmetry theorem is introduced.

# Chapter VI

# Applications of Combinatorial Species of Structure

As more sophisticated requests of the theory of combinatorial species of structure were made in order to provide explanations and generalisations of algebraic identities of power series such as Lagrange inversion, the theory extended its scope of ideas and thus the range of applications. The extensions presented in this chapter provide consistent combinatorial explanations and understanding for a greater range of algebraic operations, which are present in the manipulations of cluster and virial expansions.

The theorems and key concepts in this section come from Bergeron Labelle and Leroux [BLL98]. Further details have been added to give further explanations of proofs and concepts.

## VI.1 Virtual Species

The combinatorial calculus developed so far for species of structure has a very important gap. The issue is in the absence of a convenient combinatorial operation of subtraction. This problem is resolved by introducing a concept of virtual species. The introduction of such species has many consequences. The most important to us are:

1. We obtain a combinatorial meaning for the multiplicative inverse $\frac{1}{F}$ for any species such that $|F[\emptyset]| = 1$.

2. We obtain a combinatorial meaning for the inverse of substitution $G^{(-1)}$ for suitable species $G$

3. This provides a method of realising connected components for a wide class of species of structure

**Definition.** *Let $F$ be a (weighted) species. A species $G$ is said to be a subspecies of $F$ (written $G \subset F$) if it satisfies the following two conditions:*

*i) for any finite set $U$, $G[U] \subset F[U]$ and, in the weighted case, the weighting on $G[U]$ is induced from that of $F[U]$*

*ii) for any $\sigma : U \to V$, $G[\sigma] = F[\sigma|_{G[U]}]$.*

By analogy with how $\mathbb{Z}$ is constructed from $\mathbb{N}$, we give the following definition:

**Definition.** *A Virtual Species is an element of the quotient set:*

$$Virt = (Spe \times Spe)/\sim \tag{VI.1.1}$$

*where the equivalence relation $\sim$ is defined by:*

$$(F, G) \sim (H, K) \iff F + K \simeq G + H \tag{VI.1.2}$$

*where $\simeq$ denotes species isomorphism (not equipotence).*

*One then writes:*

$$F - G = \text{ class of } (F, G) \text{ according to } \sim \tag{VI.1.3}$$

*to denote any virtual species. The pair $(F, G)$ is called a representative of $F - G$.*

**Remark 32.** *The set of virtual species is a commutative ring under addition and multiplication.*

*We note also that $(F - G)(x) = F(x) - G(x)$.*

**Definition.** *Two species of structure $F$ and $G$ are said to be unrelated if the only subspecies of $F$ which is isomorphic to a subspecies of $G$ is the empty species. A virtual species $\Phi$ is said to be written in reduced form $\Phi = \Phi^+ - \Phi^-$ if the species $\Phi^+$ and $\Phi^-$ are unrelated.*

In the paper by Stell [Ste65], we see that the intent to invert a relationship between the connected and two-connected graphs relies on taking a logarithm. The desire is to invert the exponential function and composition of functions in combinatorics. To do so on the level of species of structure adds power to the generalisability of such a method. Furthermore, from the power series expansion of $\ln(1 + A)$, we see the necessity of introducing these virtual species, since terms have minus signs.

**Example** (Compositional Inverse of the Set Species $\mathcal{E}$)**.** *In order to understand a virtual species $F$ such that $F \circ \mathcal{E} = 1$, we define $\mathcal{E}_+ = \mathcal{E} - 1$ and set:*

$$\mathcal{E}^{-1} = (1 + \mathcal{E}_+)^{-1} = \sum_{k=0}^{\infty} (-1)^k (\mathcal{E}_+)^k \tag{VI.1.4}$$

*It can be shown that:*

$$(\mathcal{E}^{-1})^+ = 1 + (\mathcal{E}_+)^2 + (\mathcal{E}_+)^4 + \cdots \tag{VI.1.5}$$

$$(\mathcal{E}^{-1})^- = (\mathcal{E}_+) + (\mathcal{E}_+)^3 + \cdots \tag{VI.1.6}$$

*Thus an $\mathcal{E}^{-1}$-structure is an ordered partition (ballot) either with an even number of classes (positive part) or odd number of classes (the negative part).*

We define the species $\mathcal{E}_+$ as:

$$\mathcal{E}_+[U] = \begin{cases} \{U\} & \text{if } U \neq \emptyset \\ \emptyset & \text{if } U = \emptyset \end{cases} \tag{VI.1.7}$$

From the fact that multiplication of like species enforces an order on the underlying set ( explained in Section V.2.2), we see that $\mathcal{E}_+^k$ is precisely an ordered $k$-partition, since we are not allowed empty sets. The interpretation of composing with log is that we make an ordered partition with sign $(-1)^k$, where $k$ is the number of sets in the partition.

**Example** (Combinatorial Logarithm)**.** *The notion of connected components of a species of structure is a very useful concept but it can often be hard to find what the 'connected components' are and to invert the relationship. A species of structure $F$, which can be written in the form:*

$$F = \mathcal{E} \circ G \tag{VI.1.8}$$

*where $G(0) = 0$, admits the notion of connected components The connected components are enumerated by the species $G$, which is denoted: $G = F^c$. The main result is:*

**Proposition VI.1.1.** *Let $F$ be a species of structures satisfying the condition $F(0) = 1$. Then there exists a unique virtual species $\Gamma$ satisfying the combinatorial equation:*

$$F = \mathcal{E}(\Gamma) \tag{VI.1.9}$$

*Using the $\mathcal{E}_+^{(-1)}$ defined in (VI.1.4), we can define $\Gamma = \mathcal{E}_+^{(-1)} \circ F_+$, where $F_+ = F - 1$.*

### VI.1.1 General Combinatorial Inverses

The canonical decomposition of a species $F$ is the series:

$$F = F_0 + F_1 + F_2 + \cdots \tag{VI.1.10}$$

where $F_k$ is the structure which for sets $U$ with $|U| = k$, we get $F_k[U] = F[U]$, but for $V$ with $|V| \neq k$, $F_k[V] = \emptyset$.

**Proposition VI.1.2.** *Let $\Psi$ be a virtual species of structure with canonical decomposition of the form:*

$$\Psi = X + \Psi_2 + \Psi_3 + \cdots \tag{VI.1.11}$$

*i.e. a species where $\Psi_0 = 0$ and $\Psi_1 = X$. Then there exists a unique virtual species $\Psi^{(-1)}$ such that*

$$\Psi^{(-1)} \circ \Psi = \Psi \circ \Psi^{(-1)} = X \tag{VI.1.12}$$

The proof is found in [BLL98]. The key element is that the inverse may be given by:

$$\Psi^{(-1)} = \sum_{k \geq 0} (-1)^k \Delta_\Psi^k(X) \tag{VI.1.13}$$

where we define the linear operator $\Delta_\Psi : \mathrm{Virt} \to \mathrm{Virt}$, by the formula:

$$\Delta_\Psi(\Phi) = \Phi \circ \Psi - \Phi \tag{VI.1.14}$$

The importance of inversion allows important connections to be made between particular structures. This is useful in Chapter VII.

## VI.2 Lagrange Inversion

Lagrange inversion provides an algebraic method to invert (formal) power series. It has been generalised to the Lagrange-Good inversion, which is conveyed in Chapter IX. Combinatorics aims to provide bijective proofs. In the theory of combinatorial species of structure we wish to have an interpretation of Lagrange inversion. This has been shown in [BLL98, GL95].

In order to identify the $z^n$ coefficient of an analytic function $f(z)$, we use

Cauchy's residue theorem and calculate the contour integral:

$$\frac{1}{2\pi i}\oint_C \frac{f(z)}{z^{n+1}}\mathrm{d}z \tag{VI.2.1}$$

where $C$ is a contour containing the origin.

The main ingredient to prove the Lagrange inversion formula in complex analysis is the change of variable formula:

$$\mathsf{Res} f(x)\mathrm{d}x = \mathsf{Res} f(w(t))w'(t)\mathrm{d}t \tag{VI.2.2}$$

where $w$ is an invertible change of variables and $f$ is a meromorphic function.

**Theorem VI.2.1** (Lagrange Inversion)**.** *For any function $R$, satisfying $R(0) = 0$ and $R'(0) \neq 0$ and two variables $x$ and $t$, related through $R$ by the equation:*

$$t(x) = xR(t(x)), \tag{VI.2.3}$$

*we have, for any power series $g(x)$ in $x$, the following identification of coefficients:*

$$[x^n]g(x) = \frac{1}{n}[t^{n-1}]g'(t)R(t)^n \tag{VI.2.4}$$

*where $g'(t) = \frac{d}{dt}g(x(t))$.*

*Proof Based on Contour Integrals.* Let $u(t) := \frac{t}{R(t)}$, then we may make the change of variables $x = u(t)$ inside the evaluation of the contour integral:

$$\begin{aligned}
[x^n]g(x) &= \mathsf{Res}\frac{g(x)}{x^{n+1}}\mathrm{d}x && \text{(VI.2.5)}\\
&= \mathsf{Res} g(t)\frac{u'(t)}{u(t)^{n+1}}\mathrm{d}t = \mathsf{Res}\left(\frac{g'(t)}{nu(t)^n} - \left(\frac{g(t)}{nu(t)^n}\right)'\right)\mathrm{d}t && \text{(VI.2.6)}\\
&= \mathsf{Res}\frac{g'(t)}{nu(t)^n}\mathrm{d}t = \frac{1}{n}\mathsf{Res}\frac{g'(t)R(t)^n}{t^n}\mathrm{d}t && \text{(VI.2.7)}\\
&= \frac{1}{n}[t^{n-1}]g'(t)R(t)^n && \text{(VI.2.8)}
\end{aligned}$$

The total derivative term in (VI.2.6) disappears, because it is analytic and so cannot contribute any residue.

□

The set of power series $f \in \mathbb{C}[[x]]$ such that $f(0) = 0 \neq f'(0)$ constitutes a group under the operation of substitution. Using the Lagrange inversion formula,

one can write explicitly the coefficients of the inverse formal power series $f^{-1}(x)$ of $f(x)$ as follows:

$$f^{-1}(x) = \sum_{n\geq 1} \left(\frac{d}{dt}\right)^{n-1} \left(\frac{1}{f(t)}\right)^n \Big|_{t=0} \frac{x^n}{n!} \tag{VI.2.9}$$

If we denote by $A(x)$ the inverse of $f(x)$ for substitution and set:

$$R(x) = \frac{x}{f(x)} \in \mathbb{C}[[x]] \tag{VI.2.10}$$

The series $A(x)$ is determined by the functional equation:

$$A(x) = xR(A(x)) \tag{VI.2.11}$$

More generally, for any power series $F(x)$, if

$$F(A(x)) = \sum_{n\geq 0} b_n \frac{x^n}{n!} \tag{VI.2.12}$$

then we have $b_0 = F(0)$ and for any $n \geq 1$

$$b_n = \left(\frac{d}{dt}\right)^{n-1} F'(t)(R(t))^n \Big|_{t=0} \tag{VI.2.13}$$

The aim is to generalise this relationship for $R$, a species of structure. We explicitly construct the species $\mathcal{A}_R$ of $R$-enriched rooted trees, which is the solution to the functional equation:

$$\mathcal{A}_R = X \star R(\mathcal{A}_R) \tag{VI.2.14}$$

**Definition.** *An $R$-enriched rooted tree on a finite set $U$ is the sequence $(\alpha, (R_\alpha(u))_{u\in U})$, where:*

*i) $\alpha$ is an arbitrary rooted tree on $U$*

*ii) $R_\alpha(u)$ is an $R$-structure on the fibre of the vertex $u \in U$, in this rooted tree.*

**Definition.** *The fibre of a vertex $u$ is the set $\alpha^{-1}(u)$ of immediate predecessors of $u$, when all edges of the rooted tree are oriented towards the root. The indegree of $u$, denoted by $d^-(u)$, is, by definition, the cardinality of its fibre: $d^-(u) = |\alpha^{-1}(u)|$.*

**Remark 33.** *Note that we need to be in the case $R[\emptyset] \neq \emptyset$.*

One method, which is not recommended, is to compute the coefficients re-

cursively from the functional form:

$$A_R(x) = xR(A_R(x)) \tag{VI.2.15}$$

of course $a_0 = 0$ and $a_1 = r_0$, when we define the generating functions as:

$$R(x) := \sum_{n=0}^{\infty} r_n \frac{x^n}{n!} \quad A_R(x) := \sum_{n=1}^{\infty} a_n \frac{x^n}{n!} \tag{VI.2.16}$$

The recursion is:

$$a_{n+1} = \sum_{\substack{k_1+\cdots+k_j=n \\ j\geq 0}} \frac{(n+1)!}{j!k_1!\cdots k_j!} r_j a_{k_1} \cdots a_{k_j} \tag{VI.2.17}$$

We would like to have more transparent expression, which is easier to use. Let us first consider the species $R^\lambda$, for $\lambda \in \mathbb{N}$:

$$R^\lambda(x) = (R(x))^\lambda := \sum_{n=0}^{\infty} r_n(\lambda) \frac{x^n}{n!} \tag{VI.2.18}$$

We know that $R^{\lambda+\mu} = R^\lambda R^\mu$ and this gives the binomial identity:

$$r_n(\lambda+\mu) = \sum_{k=1}^{n} \binom{n}{k} r_k(\lambda) r_{n-k}(\mu) \tag{VI.2.19}$$

The sequence $(r_n(\lambda))_{n\geq 0}$ is called the binomial type sequence.

We have seen in Section V.2.2 that, for $\lambda \in \mathbb{N}$, an $R^\lambda$-structure is an ordered $\lambda$-partition of a set, with an $R$-structure on each set in the partition. This corresponds to a function $f : U \to [\lambda]$, such that on each fibre $f^{-1}(k)$, we have an $R$-structure.

**Proposition VI.2.2** (Relationship of $a_n$ to $r_n(\lambda)$)**.** *The number $a_n$ of $R$-enriched rooted trees on a set of $n$ elements is given by:*

$$a_n = r_{n-1}(n) \tag{VI.2.20}$$

*For any $k \geq 0$, the number $a_n^{\{k\}}$ of forests of $k$ $R$-enriched rooted trees on a set of $n$ elements is given by:*

$$a_n^{\{k\}} = \frac{k}{n} \binom{n}{k} r_{n-k}(n) \tag{VI.2.21}$$

**Definition** (Endofunctions)**.** *We understand the species of Endofunctions on a fi-*

*nite set $U$ as being the set of $\gamma \subset U \times U$, where for each $x \in U$, we have a unique $y \in U$ such that $(x, y) \in \gamma$. The digraph we create from this is called the functional digraph or Sagittal Graph of the Endofunction $\psi$.*

*The transport of End$[\sigma]$ : End$[U] \to$ End$[V]$ for the bijection $\sigma : U \to V$ is End $[\sigma](\psi) = \sigma \circ \psi \circ \sigma^{-1}$ $\forall \psi \in$ End$[U]$.*

In Figure VI.1, we have the sagittal graph for the endomorphism on [11]. The arrows indicate where the endofunction takes the number. For example, the arrow from 2 to 1 indicates $\psi(2) = 1$.

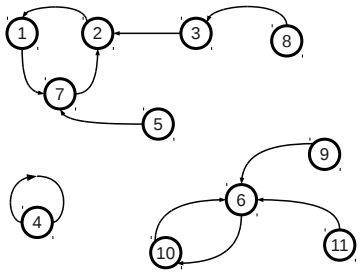



Figure VI.1: The Sagittal Graph of an Endomorphism

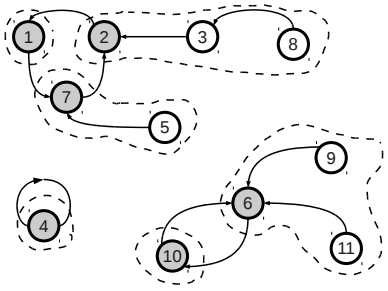



Figure VI.2: The Sagittal Graph as an Assembly of Rooted Trees, with a Permutation Structure between the Rooted Trees

**Definition** (Vertebrate Species)**.** *The vertebrate species $\mathfrak{v} = \mathfrak{a}^{\bullet\bullet}$, is the doubly-rooted tree species. The roots can be at the same vertex. It comprises of a tree and an identification of two (not necessarily distinct) vertices.*

**An example that indicates the distinction between equipotence and isomorphism of species of structure**

*We indicate here the equipotence of vertebrate species and Endomorphisms, providing an alternative proof of Cayley's formula for the number of trees*

The first stage is to understand the species isomorphism

$$\mathfrak{v} = \mathcal{L}_+(\mathcal{A}) \tag{VI.2.22}$$

where $\mathcal{L}_+$ is the species of non empty linear orderings. For this identification, we need to make use of the vertebral column.

**Definition** (Vertebral Column)**.** *Between any two vertices in a tree we have a unique path. In the vertebrate species the unique path between the two roots is called the vertebral column.*

The vertices in the vertebral column can be understood as the roots of trees emanating from this identified path. The path is a non-empty linear ordering on the roots of the trees. This gives (VI.2.22).

There is an equipotence $\mathcal{L}_+ \equiv \mathcal{S}_+$, since the number of linear orderings of $n$ items is the same as the number of permutations, $n!$. We therefore have $\mathfrak{v} \equiv \mathcal{S}_+(\mathcal{A})$. This is not a species isomorphism as cycle structure captures isomorphism classes of permutations but all linear orderings are isomorphic.

To make the connection with endomorphisms, realise that for the sagittal graph of an endomorphism, we can consider the cycles as the permutation from the outer $\mathcal{S}_+$-structure. We see that emanating from these cycles we have rooted trees, where we identify the roots of the trees with the vertices involved in the cycles.

This gives us that $v \equiv \mathrm{End}_+$ and so the number of doubly-rooted trees or vertebrates is $n^n$ (number of functions $f : [n] \to [n]$) and as a corollary the Cayley formula that the number of trees on $n$ vertices is $n^{n-2}$. We can see how this works on our example in Figure VI.2.

**Definition** ($R$-enriched endofunctions)**.** *An $R$-enriched partial endofunction on a set $U$, consists of a subset $V \subset U$ and a function $f : V \to U$ of which each fibre $f^{-1}(u)$ $u \in U$ is given an $R$-structure. When $V = U$ the function $f : U \to U$ is called an $R$-enriched endofunction.*

*We let $End_R$ and $End_R^{\mathcal{P}}$ denote the species of $R$-enriched endofunctions and partial endofunctions respectively.*

**Lemma VI.2.3** (Endomorphism Isomorphisms)**.** *Let $R$ be a species of structure and $U$ a finite set, with $V \subset U$ a fixed subset. Let $k = |V|$ and $n = |U|$. The number of $R$-enriched partial endofunctions with domain $V$ is equal to $r_k(n) = |R^n[k]|$ and we have the two combinatorial isomorphisms:*

$$End_R^{\mathcal{P}} = \mathcal{E}(\mathcal{A}_R) \star End_R \tag{VI.2.23}$$

$$End_R = \mathcal{S}(X \star R'(\mathcal{A}_R)) \tag{VI.2.24}$$

We give here a more detailed explanation of the proof found in [BLL98].

*Proof.* Consider repeatedly applying the partial endofunction $f : V \to U$. This decomposes $V$ into the two disjoint subsets, described by:

i) $x \in V$ such that for some $k \in \mathbb{N}$ $f^k(x) \notin V$.

ii) $W := \{y \in V$ such that for all $l \in \mathbb{N}$ $f^l(y) \in V\}$.

If we restrict $f$ to $W$, then $f|_W$ is an endofunction on $W$. Thus we have the splitting into the product of $\mathrm{End}_R$ and the structure, we have on the complement. On the set $U \setminus W$, we have $R$-enriched rooted trees, where the roots are the set $U \setminus V$. This is illustrated in Figure VI.3.

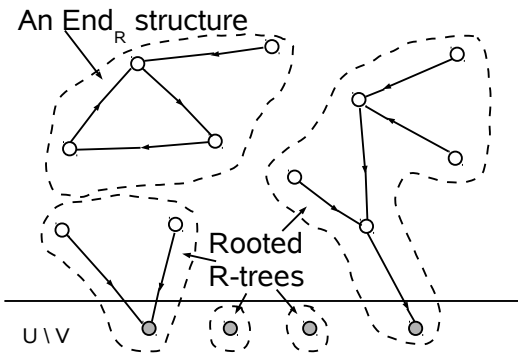



Figure VI.3: How a Partial $\mathrm{End}_R$-structure is Built Up

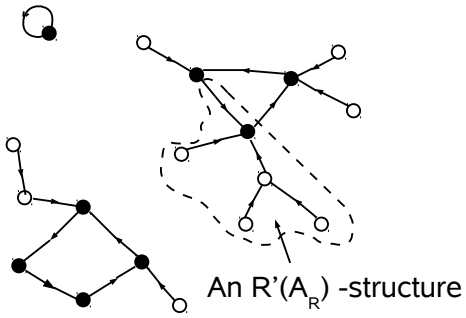



Figure VI.4: Building an Endomorphism from $\mathcal{Q}$-structures

For the second relation, we must first understand the species $\mathcal{Q} := X \star R'(\mathcal{A}_R)$.

We have a special vertex indicated by the $X$-species. The $\mathcal{A}_R$-structures are $R$-enriched rooted trees. We understand the outer $R'$-structure as incoming arrows from the roots of all the subtrees, plus an extra arrow coming from the derivative. The vertex at which these arrows point is the vertex identified by $X$. In Figure VI.4, we have marked an individual structure in the sagittal graph of an endomorphism so we can see how they are built up.

The outer permutation $\mathcal{S}$-structure gives the cycle structure between the $X$ vertices, indicating where the extra $R$-arrow comes from in each case. □

**Remark 34.** *One can imagine the $\mathcal{Q}$-structures in the language of Feynman diagrams as possessing a single 'external leg'. The cycle structure then prescribes how the legs are connected between vertices.*

This leads us to the combinatorial version of Lagrange inversion:

**Theorem VI.2.4** (Combinatorial Version of Lagrange Inversion)**.** *Let $R$ and $F$ be*

*two combinatorial species of structure.* $\forall n \geq 0$*, we have the following bijections:*

$$[i] \qquad \mathcal{A}_R^{\bullet}[n] \quad \tilde{\to} \quad (X \star R^n)[n] \tag{VI.2.25}$$

$$[ii] \qquad F(\mathcal{A}_R)^{\bullet}[n] \quad \tilde{\to} \quad (F^{\bullet} \star R^n)[n] \tag{VI.2.26}$$

$$[iii] \qquad (F(\mathcal{A}_R) \star End_R)[n] \quad \tilde{\to} \quad (F \star R^n)[n] \tag{VI.2.27}$$

The majority of the proof can be found in [BLL98].

*Proof of Theorem VI.2.4.* First note that [i] follows from [ii], by setting $F = X$. We will prove [iii] first and then [ii].

For [iii], if we consider an $(F(\mathcal{A}_R) \star \mathrm{End}_R)$-structure on the set $U = [n]$ and denote by $W$ the set of roots of the $R$-enriched rooted trees of the $F$-assembly. We decompose this whole structure as an $F$-structure on $W$ with an $R$-enriched function with domain $U \setminus W$ and codomain $U$. Since $|U| = n$, this last structure may be identified with an $R^n$-structure on $V = U \setminus W$, which gives an overall $(F \star R^n)$-structure on $U$. This is shown in Figure VI.5. This is because for each vertex in $U$, we have a pre-image in $V$. Considering the pre-images results in a partition of $V$ with some additional empty subsets for vertices that are not in the image. We have a separate $R$-structure on each of these subsets, including the empty sets, which gives the bijection with an $R^n$-structure on this subset.

For [ii], We consider a series of bijections from the $F(\mathcal{A}_R)^{\bullet}$-structure.

First of all, separate from the $F$-assembly the whole rooted tree corresponding to the distinguished point (coming from the overall rooting with the $\bullet$).

Separate the subtree rooted at this special point and join this subtree to the initial $F$-assembly.

We have now a rooted $F$-structure on this assembly and similarly to the last proof, we can understand the rest of the structure as being an $R$-enriched endofunction. This transformation is displayed in Figure VI.6. □

We can now transform these general theorems for species of structure into the classical Lagrange inversion formulæ. We start with the two functions:

$$R(x) := \sum_{n \geq 0} r_n \frac{x^n}{n!} \quad F(x) := \sum_{n \geq 0} f_n \frac{x^n}{n!} \tag{VI.2.28}$$

If $\mathcal{A}_R$ is the species of $R$-enriched rooted trees with generating function:

$$A(x) := \sum_{n \geq 0} a_n \frac{x^n}{n!} \tag{VI.2.29}$$

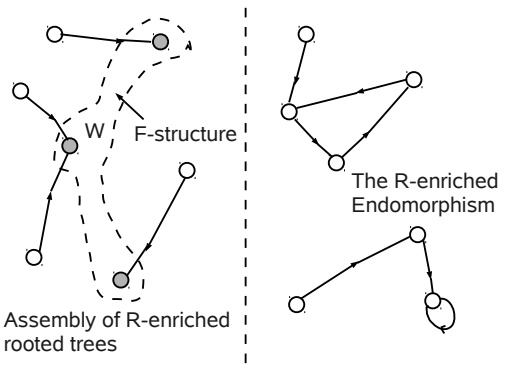



Figure VI.5: Diagram Corresponding to Species in Theorem VI.2.4 [iii]

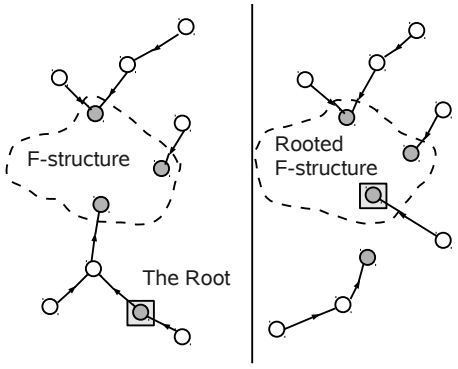



Figure VI.6: Diagram Corresponding to Species in Theorem VI.2.4 [ii]

and the relationship $A(x) = xR(A(x))$ and we write:

$$F(A(x)) := \sum_{n\geq 0} b_n \frac{x^n}{n!} \tag{VI.2.30}$$

$$F(A(x)) \exp(xR'(A(x))) := \sum_{n\geq 0} c_n \frac{x^n}{n!} \tag{VI.2.31}$$

then the three equations from our version of the Lagrange Inversion formula give us:

$$na_n = |(X \star R^n)[n]| = n|R^n[n-1]| \tag{VI.2.32}$$

$$nb_n = |(F^\bullet \star R^n)[n]| \tag{VI.2.33}$$

$$c_n = |(F \star R^n)[n]| \tag{VI.2.34}$$

We know that:

$$|G[n]| = n![x^n]G(x) = \left.\frac{d^n}{dt^n} G(t)\right|_{t=0} \tag{VI.2.35}$$

Using the definition:

$$R^n(x) := \sum_{k\geq 0} r_k(n) \frac{x^k}{k!} \tag{VI.2.36}$$

we have:

$$a_n = r_{n-1}[n] \tag{VI.2.37}$$

$$b_n = \sum_{k=1}^{n} \frac{k}{n} \binom{n}{k} f_k r_{n-k}(n) \tag{VI.2.38}$$

$$c_n = \sum_{k=0}^{n} \binom{n}{k} f_k r_{n-k}(n) \tag{VI.2.39}$$

This method of Lagrange inversion and its extension to coloured species helps us obtain an idea on the coefficients of the virial expansion, since we have

$$\rho(z) = z \frac{d}{dz} \beta P(\rho(z)) \tag{VI.2.40}$$

if we want to write pressure in terms of density. We have the two 'known' series expansions:

$$F(z) = \beta P(z) \quad \text{and} \quad R(z) = \frac{1}{\frac{d}{dz}\beta P(z)} \tag{VI.2.41}$$

We define $A(\rho)$ as the unknown inversion $z(\rho)$ and we want to calculate $F(z(\rho)) = \beta P(z(\rho))$, which will give us the power series expansion

$$\beta P(\rho) = \sum_{n=1}^{\infty} b_n \frac{\rho^n}{n!} \tag{VI.2.42}$$

We can find $b_n$ from the known coefficients by:

$$b_n = \sum_{k=1}^{n} \frac{k}{n} \binom{n}{k} f_k r_{n-k}(n) \tag{VI.2.43}$$

This gives us a method of finding the virial coefficients from the cluster coefficients $f_k$ and some operation on the cluster coefficients $r_{n-k}(n)$.

## VI.3 Dissymmetry Theorems

### VI.3.1 Statement of the Theorems

For the virial expansion, it has been shown that the coefficients are the sum over weighted two-connected graphs. for example in [MMay40] and in Section II.2. This fact can be perceived as a consequence of the *dissymmetry theorem*. The additional information the dissymmetry theorem gives us over Lagrange inversion is an interpretation of the coefficients rather than just a relationship.

**Theorem VI.3.1** (The Dissymmetry Theorem)**.** *For the species of connected graphs $\mathcal{C}$ and the species of two-connected graphs $\mathcal{B}$, we have the following relationship:*

$$\mathcal{C} + \mathcal{C}^\bullet \star \mathcal{B}'(\mathcal{C}^\bullet) = \mathcal{C}^\bullet + \mathcal{B}(\mathcal{C}^\bullet) \tag{VI.3.1}$$

Using the dissymmetry theorem to quickly obtain the virial expansion coefficients has not been done in the literature and appears to be the simplest and most powerful way to achieve the virial expansion.

This statement is true as an isomorphism of species of structure and is also true if we give the graphs weights, which factorise over blocks in our graph. The virial expansion we get from this theorem is:

$$\beta P(\rho) + \rho \sum_{n=2}^{\infty} n \frac{b_n}{n!} \rho^{n-1} = \rho + \sum_{n=2}^{\infty} \frac{b_n}{n!} \rho^n \tag{VI.3.2}$$

Which is manipulated into the form:

$$\beta P(\rho) = \rho - \sum_{n=2}^{\infty} (n-1) \frac{b_n}{n!} \rho^n \tag{VI.3.3}$$

where

$$b_n = \sum_{g \in \mathcal{B}[n]} w(g) \tag{VI.3.4}$$

Where $w$ is our weight. In the literature, the notation $\beta_n := \frac{b_{n+1}}{n!}$ is often used. In this case our virial expansion becomes:

$$\beta P(\rho) = \rho - \sum_{n \geq 1} \frac{n}{n+1} \beta_n \rho^{n+1} \tag{VI.3.5}$$

### VI.3.2 Understanding Dissymmetry Theorems

In order to understand the connected graph dissymmetry theorem, it is important to understand the general principle behind dissymmetry theorems. In a dissymmetry theorem, we have a species $\mathcal{F}$, which has a special characterising equation of the form:

$$\mathcal{F} = X \star \mathcal{R}(\mathcal{F}) \tag{VI.3.6}$$

where $\mathcal{R}$ is the relevant species of structure for this characterising relationship. This is again the idea of an $\mathcal{R}$-enriched rooted tree.

**Theorem VI.3.2** (Recursive Relationship for rooted connected graphs)**.** *We have*

*the following relationship between connected and two-connected graphs:*

$$\mathcal{C}^\bullet = X \star \mathcal{E}(\mathcal{B}'(\mathcal{C}^\bullet)) \tag{VI.3.7}$$

A simple example of this type of relationship is:

**Theorem VI.3.3** (Recursive Relationship for rooted Trees)**.** *We have the recursive relationship for trees:*

$$\mathfrak{a}^\bullet = X \star \mathcal{E}(\mathfrak{a}^\bullet) \tag{VI.3.8}$$

*Proof of Theorem VI.3.3.* Consider a rooted tree and look at its root. Consider the neighbours of the root. The removal of the root of the tree leaves a forest comprised of disjoint trees (see Figure VI.7), each of which we can perceive as being rooted by the corresponding neighbour of the root. Thus, a rooted tree can be deconstructed as: the root separately (which gives the $X$); and the decomposition of the rest of the graph into the component trees, each of which are rooted at the unique neighbour of the root, since we have only one path from each neighbour of the root to the other (i.e. through the root) they are all disconnected. This is presented by the $\mathcal{E}(\mathfrak{a}^\bullet)$-structure. We get the partition from $\mathcal{E}$ and the operation of composition and we get each rooted subtree from the $\mathfrak{a}^\bullet$. □

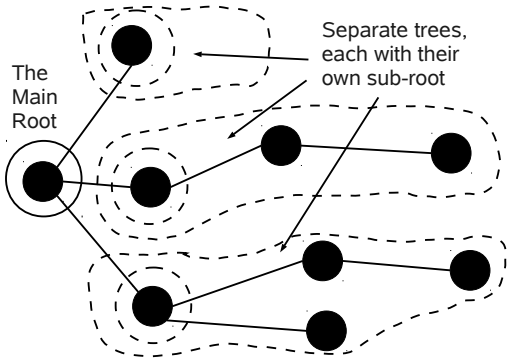



Figure VI.7: The Structure of a Rooted Tree in Theorem VI.3.3

The recursion relation for connected graphs can also be used to obtain the virial expansion, this can be found in the paper of Leroux [Ler04]. We have the expression of pressure in terms of weighted connected graphs as:

$$\beta P = \mathcal{C}_w(z) = \int_0^z \mathcal{C}'_w(t)\mathrm{d}t = \int_0^z \frac{\rho(t)}{t}\mathrm{d}t \tag{VI.3.9}$$

The relationship in Theorem VI.3.2 tells us that $t(r) = r\exp(-\mathcal{B}'_w(r))$. This means we can change our integration variable to $r$ through this relationship. The change of variables is of the form:

$$\mathrm{d}t = \left(\exp(-\mathcal{B}'_w(r)) - r\exp(-\mathcal{B}'_w(r))\mathcal{B}''_w(r)\right)\mathrm{d}r \tag{VI.3.10}$$

This gives the expression:

$$\begin{aligned}\beta P &= \int_0^\rho 1 - r\mathcal{B}''_w(r)\mathrm{d}r = \rho - \int_0^\rho r\mathcal{B}''_w(r)\mathrm{d}r \\ &= \rho - \int_0^\rho \sum_{n\geq 1} nB_{n+1}\frac{r^n}{n!}\mathrm{d}r \\ &= \rho - \sum_{n\geq 2}(n-1)B_n\frac{\rho^n}{n!}\end{aligned} \tag{VI.3.11}$$

In order to understand the recursive relationship for rooted connected graphs, we first have to understand the structure of connected graphs in terms of two-connected graphs. This has been shown as the bc-tree in Section III.2.

*Proof of Theorem VI.3.2.* First we give a mapping from a $X\star\mathcal{E}(\mathcal{B}'(\mathcal{C}^\bullet))$-structure to a rooted connected graph. We then indicate that this mapping is indeed a bijection.

Consider a $X \star \mathcal{E}(\mathcal{B}'(\mathcal{C}^\bullet))$-structure. The multiplication by the species $X$ indicates the root of the connected graph. The $\mathcal{E}$ structure partitions the non-rooted elements into subsets. Each of these subsets has a $\mathcal{B}'(\mathcal{C}^\bullet)$-structure. On each subset, we identify the $\mathcal{B}'(\mathcal{C}^\bullet)$-structure with the graph comprising of a central $\mathcal{B}'$-structure between a ghost vertex and the rooted vertices of the $\mathcal{C}^\bullet$-structures. The ghost vertices for each subset are all identified with the root vertex to construct the final connected graph, rooted at this vertex.

We need to indicate why this is a bijection and to that end we convey how each rooted connected graph can be realised as one of these structures, uniquely. The main idea is that a given point is in a given number of blocks. Each of these blocks and the rest of the graph connected to the particular block is in one of the subsets generated by $\mathcal{E}$ - this is conveyed as the unbroken sets in Figure VI.9. We get the block structure from the $\mathcal{B}'$ structure (conveyed by the graph with the light grey circles being the 'ghost' vertices and those inside squares as the other vertices) and the rest of the graph which is connected to each vertex in the block is encoded in the $\mathcal{C}^\bullet$-structure (these are the subgraphs circled with dashed lines in Figure VI.9). So

we have a unique splitting of our rooted connected graph into a structure recognised by the left hand side. □

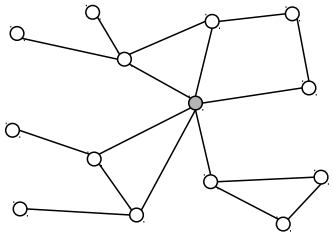


Figure VI.8: An Example of a Rooted Connected Graph

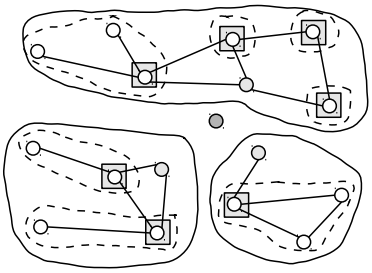


Figure VI.9: How the Graph Splits into the species in Theorem VI.3.2

### VI.3.3 Heuristics of the Dissymmetry Theorem

The overarching concept of the dissymmetry theorems is the Dissimilarity Characteristic Theorem found in [HaPa73]. This is connected to tree graphs having a well defined centre as indicated below. The first Dissymmetry Theorem is known as Otter's formula and is given below:

**Theorem VI.3.4** (Dissymmetry Theorem for Trees)**.** *Let* $\mathfrak{a}^-$ *represent the species of trees rooted at an edge and* $\mathfrak{a}^{\bullet-\circ}$ *represent the species of trees rooted at an oriented edge. Then we have the relationship:*

$$\mathfrak{a} + \mathfrak{a}^{\bullet-\circ} = \mathfrak{a}^{\bullet} + \mathfrak{a}^{-} \tag{VI.3.12}$$

**Definition.** *The eccentricity of a vertex* $v$ *in a tree* $T$ *is:* $\varepsilon(v) = \max\{d(v,w)|w \in V(T)\}$ *The radius of a tree is:* $radius(T) = \min\{\varepsilon(v)|v \in V(T)\}$ *The centre of a tree is:* $\{v \in V(T)|\varepsilon(v) = radius(V)\}$.

**Definition.** *The bc-centre is the centre of the block-cutpoint tree as shown in Figure VI.10.*

**Remark 35.** *A priori the centre of a tree may be a set of cardinality strictly greater than* 1*. We are working with a special class of trees for bc-trees, which are bipartite and importantly have all their leaves in a single set. This is sufficient to get a unique centre.*

The idea is that the vertex $w$ for which the eccentricity of a vertex is attained is necessarily a leaf. If not then we always have a neighbour away from the original vertex which allows us to increase eccentricity by 1. So any articulation point has odd eccentricity and any block has even eccentricity and so the centre can contain only vertices from one of the two sets in the bipartite graph. The centre cannot contain two points of distance at least two from each other and hence can only be a singleton, so the bc-centre is well defined as a single point. An example of the bc-centre is shown in Figure VI.10.

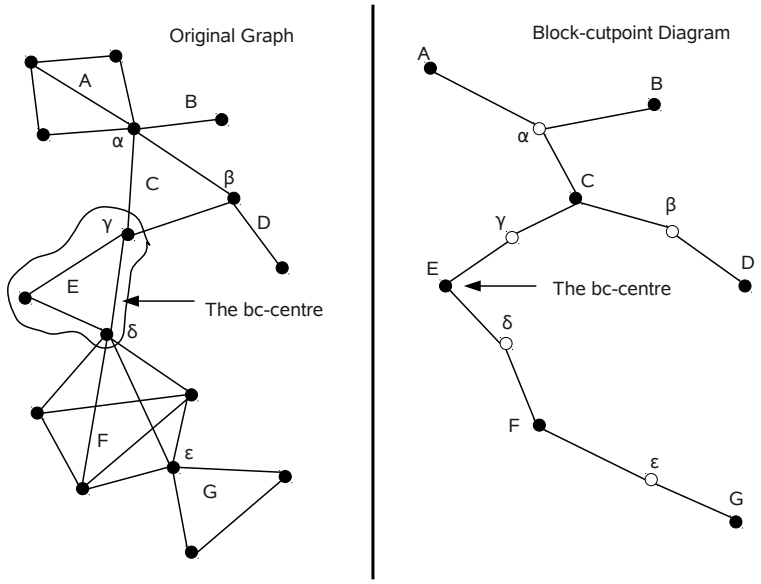



Figure VI.10: An Example of the bc-centre of a Connected Graph

We understand equation (VI.3.1) by perceiving the left hand side as the sum of a graph rooted at the bc-centre of a connected graph ($\mathcal{C}$) with the graphs rooted outside of the centre (in bijection with $\mathcal{C}^\bullet \times \mathcal{B}'(\mathcal{C}^\bullet)$). The right hand side can be viewed as graphs pointed at a vertex ($\mathcal{C}^\bullet$) or at a block ($\mathcal{B}(\mathcal{C}^\bullet)$).

### VI.3.4 Proof of the Dissymmetry Theorem

We repeat the dissymmetry theorem equation here for ease of following the proof. The relationship between connected graphs $\mathcal{C}$ and two-connected graphs $\mathcal{B}$ called the dissymmetry theorem can be written as:

$$\mathcal{C} + \mathcal{C}^\bullet \star \mathcal{B}'(\mathcal{C}^\bullet) = \mathcal{C}^\bullet + \mathcal{B}(\mathcal{C}^\bullet) \tag{VI.3.13}$$

*First Proof of VI.3.1.* We realise that the right hand side of (VI.3.13) can be understood as a connected graph rooted at a vertex, $\mathcal{C}^\bullet$, or rooted at a block in $\mathcal{B}(\mathcal{C}^\bullet)$. The latter statement can be seen by relating the composite structure to the connected graph with the outer $\mathcal{B}$-structure as the rooted block. The rooted connected graphs emanate from each vertex of this block and the vertices in the identified block are the roots of the inner rooted connected graph structures.

We now have to show that the left hand side can be understood as connected graphs rooted at either a vertex or a block.

Consider the species:

$$\mathcal{C}^\bullet \star \mathcal{B}'(\mathcal{C}^\bullet) \tag{VI.3.14}$$

First of all, we understand how this can be understood in terms of a connected graph. We have individual rooted connected graphs, which are connected together by the outer two-connected graph structure. The outer two-connected structure is understood to be between the roots of the corresponding connected graphs and the ghost vertex is identified with the root of the rooted connected structure it is multiplied by. We see that this gives a connected graph where a block (the outer two-connected graph) and a vertex (the root of the rooted connected graph structure that multiplies the $\mathcal{B}'(\mathcal{C}^\bullet)$-structure.)

Therefore, this is a connected graph rooted at a pair $(B, v)$, where $B$ is a block and $v$ is a vertex. There are two cases:

i) The vertex is internal to a block, that is, it appears in only one block. In this case the block label in the pair is redundant and this can simply be mapped to the $\mathcal{C}^\bullet$-structure rooted at this vertex.

ii) The vertex is an articulation point, which necessarily appears in multiple blocks. In this case we may consider the digraph corresponding to the bc-tree, where every edge is oriented away from the centre. This can be found in Chapter III. In this case, since any vertex has precisely one incoming vertex, excepting the centre, then for each edge we may uniquely associate the vertex it points towards in this oriented version. Therefore, we map the corresponding graph to the block or vertex rooted connected graph corresponding to the vertex the edge points at in the bc-tree. We realise that every vertex in the bc-tree is covered by this except for the centre. The centre precisely provides us with the missing object to root at.

The individual connected graph on the left hand side can thus be canonically interpreted as being rooted at the bc-centre of the connected graph.

We therefore have a bijection between the two sides of (VI.3.13), which is independent of the labelling of the graph and thus a combinatorial isomorphism. □

**Remark 36.** *A more 'constructive' proof looking at how the connected graphs relate to each of the structures in an enumerative way is given in the next subsection. This was done before being aware of the power given to the above interpretation and has similarities.*

### VI.3.5 A More Direct Approach

Firstly, there are two lemmas on the species involved in the relationship:

**Lemma VI.3.5.** *For a given graph $c$ the number of ways of giving it a $\mathcal{B}(\mathcal{C}^\bullet)$-structure is $n_B(c)$ = the number of blocks in the graph $c$.*

*Proof.* Firstly, consider the $\mathcal{B}(\mathcal{C}^\bullet)$-structure and how it corresponds to a connected graph. We have first a partition and on each set in the partition a $\mathcal{C}^\bullet$-structure, which is a connected graph where one vertex is identified. It is natural to consider the $\mathcal{B}$-structure between the sets in the sets as picking the 'pointed' vertex in each set in order to make a $\mathcal{B}$-subgraph.

**Claim.** *Every 'pointed vertex' is either an articulation point in the graph or in a singleton set in the partition and thus internal to the block created by the $\mathcal{B}$-structure.*

If the point is in a set of cardinality strictly greater than one, then the removal of the pointed vertex would mean that the rest of the points in the set would be disconnected from any point not in this set. We cannot have the whole set as a valid partition as we do not have a two-connected graph on a single point, hence it will always be an articulation point.

This therefore means that the overall $\mathcal{B}$-structure must point at a particular block in the graph $c$. Hence the number of ways of creating a graph with a $\mathcal{B}(\mathcal{C}^\bullet)$-structure is precisely $n_B(c)$. □

**Lemma VI.3.6.** *For a given graph $c$, the number of ways of giving it a $\mathcal{C}^\bullet \star \mathcal{B}'(\mathcal{C}^\bullet)$-structure is $v(c)+\sum_A(B_A-1)$, where $v(c)$ is the number of vertices in $c$ and $B_A$ is the number of blocks an articulation point appears in. The sum is over all articulation points in $c$.*

*Proof.* As in the previous proof we have a structure for each block, but within each block we can point at one of the subsets. It makes sense to perceive that the pointed vertex overall is the pointed vertex within the pointed set. Every point that isn't

an articulation point will get pointed precisely once as it will appear in a singleton set when the block is chosen as the pointed block. Then when we consider the articulation point we realise that it can be pointed precisely the number of times it appears in a block. Thus we achieve the given formula. □

*Alternative Proof of VI.3.1.* This is shown by partitioning the set of each species of structure according to connected graphs. From VI.3.5 we have $n_B(c)$ such graphs from the $\mathcal{B}(\mathcal{C}^\bullet)$ term. We have $v(c)$ from the $\mathcal{C}^\bullet$ term. We have 1 from the $\mathcal{C}$ term. From VI.3.6 we have $v(c) + \sum_A (B_A - 1)$ from the $\mathcal{C}^\bullet \star \mathcal{B}'(\mathcal{C}^\bullet)$ term. Thus we are left to show that:

$$n_B(c) - 1 = \sum_A (B_A - 1) \tag{VI.3.15}$$

This comes from considering the block-cutpoint tree. Starting from a block leaf and taking all the articulation points connected to this (just one) in the first layer and when counting $B_A - 1$ we leave out this leaf block and get all blocks distance two away from the leaf. We then consider the articulation points distance three away from the leaf. We count the $B_A - 1$ by ignoring the blocks distance two away and continue doing this. We don't run into any problems since all leaves are blocks and it is a tree so acyclic. We only miss out the initial block and so we get the required result. □

## VI.4 Coloured Species

A generalisation of labelled species is coloured species. Instead of having a set as the basis for the structure, we have a coloured set. Although the objects in the theory of species of structure already have labels, we may want to group them under particular types. Faris in [Far10], gives the application of a colour denoting particle location in a statistical mechanical application. However, in Chapter IX, it is also used to denote the type of particle, that is whether, for example, it is a carbon dioxide or oxygen molecule in a mixture of gases. The name 'coloured' for this concept arises from the subject of graph colourings.

**Definition** (Coloured Species of Structure)**.** *For the underlying set $U$, we have a fixed palette of colours $\mathcal{Q}$ and a function $a : U \to \mathcal{Q}$, which gives a colour to each member of the underlying set. The main change is to what the isomorphisms are for a coloured species of structure.*

*For two coloured sets $U$ and $V$, with respective colourings $a : U \to \mathcal{Q}$ and $b : V \to \mathcal{Q}$, a bijection $\phi : U \to V$ is an isomorphism if it preserves the colours, that*

*is:* $b \circ \phi = a$.

For the Kotecký-Preiss condition, this notion is used with the special case of graphs with edge weights:

$$w(E_{i,j}) = t(a(i), a(j)) \tag{VI.4.1}$$

where $a$ is the colour-function determining the location of the labelled particle and $t$ is the appropriate Mayer function, which depends on the location of the two particles.

Often with coloured sets, all that matters is how many labels have a particular colour and this motivates the use of multi-indices:

**Definition** (Multi-indices)**.** *A multi-index is a function* $N : \mathcal{Q} \to \mathbb{N}_0$*, giving each colour a non-negative integer value, which corresponds to the number of the given colour we have.*

There are two important pieces of notation used with the multi-index:

$$\text{Order} \quad n = |N| = \sum_{p \in \mathcal{P}} N(p) \qquad \text{and} \qquad \text{Factorial} \quad N! = \prod_{p \in \mathcal{P}} N(p)! \tag{VI.4.2}$$

The number of coloured sets corresponding to a given multi-index is the multinomial coefficient: $\frac{n!}{N!}$.

With coloured species of structure there are two forms of composition:

The scalar composition is just the same as the composition of species. This is when the outer structure $F$ in $F \circ G$ does not depend on the palette of colours $\mathcal{Q}$.

For 'coloured composition', we can have the outer $F$-structure depending on a palette $\mathcal{Q}$ of colours and we have a collection of $G_q$-structures, indexed by $q \in \mathcal{Q}$, so that on our set $U$, we have a partition into subsets $V_1, \cdots V_n$, each with a corresponding $G_{q_i}$-structure on $V_i$. The outer $F$-structure will be an $F$-structure on the coloured set $a : [n] \to \mathcal{Q}$ where $i \mapsto q_i$. This idea is very important with the multispecies expansions in Chapter IX.

## VI.5 Combinatorial Interpretation of the Kotecký-Preiss Criterion

We recall the Kotecký-Preiss Criterion for convergence of cluster expansions (from Section I.4, Theorem I.4.1):

If, for a polymer set $\mathfrak{P}$, with incompatibility relation $\iota$, we have the inequality:

$$\sum_{\gamma'|\gamma'\iota\gamma} e^{a(\gamma')+d(\gamma')} \left|\Phi(\gamma')\right| \leq a(\gamma) \tag{VI.5.1}$$

where $a, d : \mathfrak{P} \to \mathbb{R}$ are positive functions and $\Phi : \Gamma \to \mathbb{C}$ is the polymer functional, then we have the following inequality:

$$\sum_{\substack{C\in\mathcal{B}\\ C\iota\gamma}} \left|\Phi^T(C)\right| e^{a(C)} \leq a(\gamma) \tag{VI.5.2}$$

where $\mathcal{B}$ denotes the family of finite subsets of $\Gamma$ and we write $C\iota\gamma$ if $\exists\gamma' \in C$ such that $\gamma'\iota\gamma$. We have the relationship for $\Phi^T$:

$$\Phi^T(C) = \sum_{B|B\subset C} (-1)^{|C|-|B|} \log \mathcal{Z}(B, \Phi) \tag{VI.5.3}$$

where:

$$\mathcal{Z}(B, \Phi) := {\sum_{A\subset\mathfrak{P}}}' \prod_{\gamma\in A} \Phi(\gamma) \tag{VI.5.4}$$

where the sum is over all subsets where the polymers are mutually compatible.

To make the connection with coloured graphs which have the edge weight $t(p, q)$, we state here how the terms correspond to those that are used in the fixed point ideas.

The relationship to coloured structures, is that $\Gamma$ represents the 'colour' set and a polymer $\gamma$ is a colour. The edge weight $t(p, q)$ is such that:

$$t(p, q) = \begin{cases} -1 & \text{if } p\,\iota\,q \\ 0 & \text{otherwise} \end{cases} \tag{VI.5.5}$$

The function $d$ is included in the Kotecký-Preiss condition to deal with a normalisation that is required when our set of polymers is large, but doesn't feature in Faris' interpretation of this.

Let $\mathfrak{a}_p^\bullet$ be the species of rooted trees with root of colour $p$. We let $X_p$ be the one-point coloured set indicator species for the designated colour $p$ , with weight 1. We let $\hat{\mathcal{E}}^p = \mathcal{E} \circ \mathcal{E}'_{2,p}$ be the edges to coloured set indicator species that assigns to a coloured set $a$ the product $\prod_j t(p, a(j))$. The fixed point rooted tree equation can be cast in the form (a generalisation of the notion of rooted $R$-enriched tree structure

covered in Section VI.2):

$$\mathfrak{a}_p^\bullet = X_p \star (\hat{\mathcal{E}}^p \circ \mathfrak{a}^\bullet) \tag{VI.5.6}$$

This gives us an exponential generating function of the form:

$$\mathfrak{a}_p^\bullet(w) = w_p \exp\left(\sum_q t(p,q)\mathfrak{a}_q^\bullet(w)\right) \tag{VI.5.7}$$

If we set $\mathfrak{a}_p^\bullet = z_p$, we have the fixed point equation:

$$z_p = w_p \exp\left(\sum_q t(p,q) z_q\right) \tag{VI.5.8}$$

We have an equation in the form $\phi(z) = z$. It has a unique formal power series solution in which $z$ is expressed as a formal power series in $w$. Furthermore, the solution is given by iteration starting with 0. This follows in the way described for $R$-enriched trees in Section VI.2.

We let $z^{(0)} = 0$ and define the sequence $z^{(k)}$ inductively by $z^{(k+1)} = \phi(z^{(k)})$. Then for $k \geq 1$, the term $z^{(k)}$ gives the contribution of all trees of depth $\leq k-1$. Furthermore $z^{(k)}$ converges to $z$ in the sense that each coefficient in the power series of $z$ is achieved at some $k$. To see this, consider a tree on $n$ points. Then the depth of the tree is at most $n-1$. If $k \geq n$, the term $z^{(k)}$ contains the contribution of all trees of depth at most $n-1$ and so includes the contribution of this particular tree. For $k \geq n$ the term $z^{(k)}$ contains all contributions of trees on $n$ veritces.

We take $t(p,q) \geq 0$ and $w_p \geq 0$.

**Theorem VI.5.1** (Kotecký-Preiss Condition in Combinatorics)**.** *Consider the tree exponential generating function with colour pair weight factors $t(p,q) \geq 0$, and take the variables $w_p \geq 0$. Suppose the Kotecký-Preiss condition is satisfied, that is there exists a finite vector $\mathbf{x} \geq 0$ such that*

$$w_p \exp\left(\sum_q |t(p,q)| x_q\right) \leq x_p \tag{VI.5.9}$$

*Then the power series expansion of $\mathfrak{a}_p^\bullet(w)$ converges for the given $w$ and has absolute value bounded by $x_p$.*

**Remark 37** (Knaster-Tarski Fixed Point Theorem)**.** *The natural setting for such a theorem is on a complete lattice $L$. This is a partially ordered set for which each subset has an infimum and a supremum. We have a function $\phi : L \to L$, which is increasing in the sense that $x' \leq x''$ implies $\phi(x') \leq \phi(x'')$. The Knaster-Tarski*

*Theorem (see [Tar55]) says that an increasing function $\phi$ from a complete lattice to itself always has a fixed point. In fact it has a least fixed point $z = \inf\{y|\phi(y) \leq y\}$.*

The corresponding lattice fo this case is the set of positive vectors indexed by colour $L = [0, +\infty]^{\mathcal{P}}$.

The proof of this is relatively straightforward. We define the set:

$$S := \{y \mid \phi(y) \leq y\} \tag{VI.5.10}$$

Since $L$ is a complete lattice, we have inf $S$, call it $z$. For any $y \in S$, $z \leq y$ and since $\phi$ is increasing $\phi(z) \leq \phi(y) \leq y$ and so $\phi(z)$ is a lower bound for $S$. $z$ is the greatest lower bound and so $\phi(z) \leq z$. $\phi$ is increasing and so $\phi(\phi(z)) \leq \phi(z)$, which means $\phi(z) \in S$, so $\phi(z) \leq z$ and hence $\phi(z) = z$.

We can solve the fixed point equation by iteration, starting with the least element of $L$. If we denote by $u^{(k)}$ the $k$th iterate. If $u^{(k-1)} \leq u^{(k)}$, then by the fact $\phi$ is increasing $\phi(u^{(k-1)}) \leq \phi(u^{(k)})$ and so $u^{(k)} \leq u^{(k+1)}$. We know $u^{(1)} = \phi(u^{(0)}) \geq u^{(0)}$, since $u^{(0)}$ is least element and so we have the base case and hence that $u^{(k)}$ is an increasing sequence by induction. We note that $u^{(k)} \leq z$, the fixed point and so if we let $z' = \sup_k u^{(k)}$, then $z' \leq z$, but we have monotone convergence property, which means $u^{(k)}$ converges and that $z' = \phi(z')$ and so $z' = z$, since $z$ is the least fixed point.

**Theorem VI.5.2** (Cluster Expansion Convergence from Kotecký-Preiss)**.** *Consider the equilibrium gas system with interactions $-1 \leq t(p, q) \leq 0$ and activities $w_p$ satisfying $|w_p| \leq w_p^\star$. Suppose there are finite $x_p \geq 0$ such that the Kotecký-Preiss condition*

$$w_p^\star \exp\left(\sum_q |t(p, q)|x_q\right) \leq x_p \tag{VI.5.11}$$

*is satisfied, so that the corresponding rooted tree series $\mathfrak{a}_p^\bullet(w^\star)$ (defined with weights $0 \leq |t(p, q)| \leq 1$) converges to a value bounded by $x_p$. Then the series $\mathcal{C}_p^\bullet(w)$ for the expected number of particles at a site $p$ converges absolutely for $|w| \leq w^\star$, and its value satisfies:*

$$|\mathcal{C}_p^\bullet(w)| \leq x_p < +\infty \tag{VI.5.12}$$

If we consider only the pinned connected function or derivative of the connected function and make the change of variables: $x_p = w_p^\star e^{a_p}$, we have an alternative version.

**Theorem VI.5.3** (Cluster Expansion for pinned connected functions)**.** *Consider the equilibrium gas system with interactions $-1 \leq t(p, q) \leq 0$ and activities $w_p$,*

*satisfying* $|w_p| \leq w_p^\star$. *Suppose there are finite* $a_p \geq 0$, *such that the Kotecký -Preiss condition*

$$\sum_q |t(p,q)| w_q^\star e^{a_q} \leq a_p \tag{VI.5.13}$$

*is satisfied. Then the pinned connected function series* $\frac{\partial \mathcal{C}(w)}{\partial w_p}$ *converges absolutely, and its value satisfies:*

$$\left| \frac{\partial \mathcal{C}(w)}{\partial w_p} \right| \leq e^{a_p} < +\infty \tag{VI.5.14}$$

To stress the link with the original Kotecký-Preiss condition, if we make the substitutions/translations in the condition (VI.5.13):

$$|t(p,q)| = \begin{cases} 1 & \text{if } p \iota q \\ 0 & \text{otherwise} \end{cases} \qquad p = \gamma \quad q = \gamma' \quad a_p = a(p) = a(\gamma) \quad \text{and } w_q = |\Phi(\gamma')| \tag{VI.5.15}$$

we get the condition:

$$\sum_{\gamma' | \gamma' \iota \gamma} |\Phi(\gamma')| e^{a(\gamma')} \leq a(\gamma) \tag{VI.5.16}$$

which is precisely the Kotecký-Preiss conditions if we neglect the $d$-function, present for normalisation.

The link between these two is to observe if we substitute $w_p^\star e^{a_p}$ for $x_p$ in (VI.5.11), then we get:

$$w_p^\star \exp\left( \sum_q |t(p,q)| w_p^\star e^{a_q} \right) \leq w_p^\star e^{a_p} \tag{VI.5.17}$$

We can cancel the $w_p^\star$s and take logarithms, since ln is an increasing function, and get:

$$\sum_q |t(p,q)| w_q^\star e^{a_q} \leq a_p \tag{VI.5.18}$$

which is (VI.5.13). We know this gives us conclusion (VI.5.12) and for pinned graphs we have to manipulate:

$$|w_p| \left| \frac{\partial \mathcal{C}(w)}{\partial w_p} \right| \leq w_p^\star e^{a_p} \tag{VI.5.19}$$

since $|w_p| \leq w_p^\star$, we have a stronger condition for the derived connected function than the rooted connected function.

## VI.6 Cluster Expansion Convergence via Tree Fixed Point

Suppose there are finite $x_p \geq 0$ so that the Kotecký-Preiss condition:

$$w_p \exp\left(\sum_q |t(p,q)| x_q\right) \leq x_q \tag{VI.6.1}$$

holds. Then the sum $\mathcal{C}^\bullet_p(w)$ converges absolutely and in fact:

$$0 \leq \mathcal{C}^\bullet_p(w) \leq -\mathcal{C}^\bullet_p(-w) \leq x_p < \infty \tag{VI.6.2}$$

In the paper by Ueltschi [Uel04], an inductive approach is taken towards obtaining bounds similar to those of Kotecký-Preiss. The inductive bound relies on the observation that when we remove a vertex from a connected graph, we split the graph into connected components with each component having at least one edge to the removed point. The way to express such an induction in the language of combinatorial structures is through the combinatorial identity:

$$\mathcal{C}^\bullet_p = X_p \star (\mathcal{E} \circ (\mathcal{P}^p_+ \times \mathcal{C})) \tag{VI.6.3}$$

where $\mathcal{C}^\bullet_p$ is a connected graph rooted at a point of colour $p$, $X_p$ is the one-point indicator species of colour $p$ and $\mathcal{P}^p_+$ is the species which for every finite set takes the collection of non-empty subsets and assigns an edge weight to each edge from $p$ to a vertex in the subset. We understand this formula from observing that $X_p$ represents the root and the $\mathcal{E}$-structure is the splitting into the connected components obtained when removing the root. On each of the connected components we have both $\mathcal{P}^p_+$ and $\mathcal{C}$-structures, which is why the Hadamard product is used. We have the choice of a non empty subset to be the vertices connected directly to the root and the connected graph structure on all of the vertices in each connected component.

To understand how this translates into generating series we have to build it up: The main difficulty is with the $\mathcal{P}^p_+ \times \mathcal{C}$-structure. First of all the generating function for $\mathcal{P}^p_+$ is:

$$P^p_+(w) = \sum_{N \geq 0} \frac{1}{N!} \left((1+t_p)^N - 1\right) w^N \tag{VI.6.4}$$

If we write:

$$C(w) = \sum_{N \geq 0} c_N \frac{w^N}{N!} \tag{VI.6.5}$$

then, from the action of the Cartesian product on generating functions, we have:

$$(P_+^p \times C)(w) = \sum_{N \geq 0} c_N \frac{w^N}{N!} \left( (1+t_p)^N - 1 \right) = C((1+t_p)w) - C(w) \tag{VI.6.6}$$

We thus see we have the fixed point relation:

$$\begin{aligned} C_p^\bullet(w) &= w_p \exp(C((1+t_p)w) - C(w)) \\ &= w_p \exp\left( \sum_q t(p,q) \int_0^1 C_q^\bullet((1+st_p)w) \mathrm{d}s \right) \end{aligned} \tag{VI.6.7}$$

The important assumption made in the paper [Uel04] is that $|1+t(p,q)| \leq 1$. In this case the function is simplified to put everything into an extended Kotecký-Preiss condition, since the integral on the right hand side can be bounded above by $|C_q^\bullet(w)|$ and we have an inequality for this positive quantity, which gives us precisely the fact that the function

$$w_p \exp\left( \sum_q t(p,q) \int_0^1 C_q^\bullet((1+st_p)w) \mathrm{d}s \right) \tag{VI.6.8}$$

is bounded in modulus by the tree function we have earlier.

## VI.7 The Algebraic Approach of Ruelle

In introducing the Kirkwood Salsburg equation in Section I.3, we used the algebraic approach of Ruelle in [Rue69]. The key operations in this algebraic approach mirror those which are used in combinatorial species of structure. In this section, we emphasise how this relationship holds more precisely and a combinatorial interpretation of this.

The focus is on sequences: $\boldsymbol{\psi} = (\psi(x)_n)_{n \geq 0}$, which can be viewed as coefficients of the generating function:

$$\Psi(z) = \sum_{n=0}^{\infty} \frac{\psi(x)_n}{n!} z^n \tag{VI.7.1}$$

and thus interpreted as weighted species of structure. The notation used in the book

is $\langle \boldsymbol{\chi}_\Lambda, \boldsymbol{\psi} \rangle(z)$ for the generating function, which corresponds to integrated weights:

$$\langle \boldsymbol{\chi}_\Lambda, \boldsymbol{\psi} \rangle(z) := \sum_{n=0}^{\infty} \frac{z^n}{n!} \prod_{i=1}^{n} \left( \int \chi_\Lambda(x_i) \mathrm{d}x_i \right) \psi(x)_n \tag{VI.7.2}$$

The exponential function, denoted $\Gamma$ is introduced, which is just the same as composition with the SET species $\mathcal{E}$ and is understood in the way described at the end of section V.2.2. The derivative has some added information and is written $D_X$, where $X$ is a set of points. The $X$ gives precise colours or labels for the 'ghost vertices' in this case.

We have the Boltzmann factors:

$$\psi(x)_n := \prod_{1 \leq i < j \leq n} (1 + (\exp(-\beta \Phi(x_i - x_j)) - 1)) \tag{VI.7.3}$$

which correspond to simple graphs with weight function:

$$w(g) = \prod_{\{i,j\} \in E(g)} (\exp(-\beta \Phi(x_i - x_j)) - 1) \tag{VI.7.4}$$

There is an alternative notation used in the paper by Poghosyan and Ueltschi [PoUe09]. We let $\zeta_{i,j} = \exp(\beta \Phi(x_i - x_j)) - 1$ and use the set $I$ to represent the collection of coordinates $(x_i)_{i \in I}$ and we write:

$$\Psi[I] := \sum_{G \in \mathcal{G}[I]} \prod_{\{i,j\} \in G} \zeta_{i,j} \tag{VI.7.5}$$

$$\Psi^{\star(-1)}[I] = \frac{1}{\sum\limits_{G \in \mathcal{G}[I]} \prod\limits_{\{i,j\} \in G} \zeta^\star_{i,j}} \tag{VI.7.6}$$

There are also the Ursell functions $\boldsymbol{\phi}$, defined by $\boldsymbol{\psi} = \Gamma \boldsymbol{\phi}$,(which indicates their interpretation as connected graph function), where:

$$\phi(x)_n = \sum_{g \in \mathcal{C}[n]} \prod_{\{i,j\} \in E(g)} (\exp(-\beta \Phi(x_i - x_i)) - 1) \tag{VI.7.7}$$

The grand canonical partition function and the pressure are thus written as:

$$\Xi(\Lambda, z, \beta) = \langle \boldsymbol{\chi}_\Lambda, \boldsymbol{\psi} \rangle(z) \tag{VI.7.8}$$

$$\beta V(\Lambda) P(\Lambda, z, \beta) = \langle \boldsymbol{\chi}_\Lambda, \boldsymbol{\phi} \rangle(z) \tag{VI.7.9}$$

We have the series:

$$\psi^z_\Lambda(X) := \Xi^{-1}\langle \boldsymbol{\chi}_\Lambda, D_X \boldsymbol{\psi}\rangle(z) \tag{VI.7.10}$$

$$= \frac{1}{\mathcal{G}[x]}\mathcal{G}[X+x] \tag{VI.7.11}$$

where the last line is an abuse of notation. The small $x$ indicates we have integrals over the various variables we can have in the sequence, whereas the large $X$ denotes the 'external points'. We also define the sequence:

$$\tilde{\boldsymbol{\phi}}_X := \boldsymbol{\psi}^{-1} \star D_X \boldsymbol{\psi} \tag{VI.7.12}$$

The inverse is with respect to the star multiplication. For the single point version this simplifies somewhat to:

$$\tilde{\boldsymbol{\phi}}_{(x_1)} = \boldsymbol{\psi}^{-1} \star (D_{x_1}\Gamma\boldsymbol{\phi}) = \boldsymbol{\psi}^{-1} \star D_{x_1}\boldsymbol{\phi} \star \boldsymbol{\psi} = D_{x_1}\boldsymbol{\psi} \tag{VI.7.13}$$

In general, we have:

$$\begin{aligned}\psi^z_\Lambda(X) &= \Xi^{-1}\langle \boldsymbol{\chi}_\Lambda, \boldsymbol{\psi} \star (\boldsymbol{\psi}^{-1} \star D_X\boldsymbol{\psi})\rangle(z) \\ &= \Xi^{-1}\langle \boldsymbol{\chi}_\lambda, \boldsymbol{\psi}\rangle(z) \times \langle \boldsymbol{\chi}_\lambda, \boldsymbol{\psi}^{-1} \star D_X\boldsymbol{\psi}\rangle(z) \\ &= \langle \boldsymbol{\chi}_\Lambda, \tilde{\boldsymbol{\phi}}_X\rangle(z)\end{aligned} \tag{VI.7.14}$$

We may interpret the sequence $\tilde{\boldsymbol{\phi}}_X$ as:

$$\tilde{\boldsymbol{\phi}}_X = \left(\frac{\mathcal{G}[X+(x)_n]}{\mathcal{G}[X]}\right)_{n\geq 0} \tag{VI.7.15}$$

In order to capture the inductive nature of the Kirkwood-Salsburg equations, Poghosyan and Ueltschi, within this algebraic approach, define the function (for $I \cap J = \emptyset$:

$$g(I,J) := (\Psi^{\star(-1)} \star D_I\Psi)[J] \tag{VI.7.16}$$

This satisfies the inductive relation:

$$g(I,J) = (\prod_{i\in I'}(1+\zeta_{i,\tau(I)}))\sum_{K\subset J}(\prod_{i\in K}\zeta_{i,\tau(I)})g(I'\cup K, J\setminus K) \tag{VI.7.17}$$

for $i \neq \emptyset$ and where $\tau(I)$ is a specified element of $I$ and $I' = I \setminus \{\tau(I)\}$. We have the boundary condition $g(\emptyset, J) = \delta_{\emptyset,J}$

We may view $g$ as a function $g : \mathcal{P}[n]^{(2)\prime} \to \mathbb{C}$, where the domain is under-

stood as the set of all ordered pairs of subsets of $[n]$, with the restriction that the two sets have empty intersection. If we call such a structure $\tilde{\mathcal{G}}$, then we can view $g$ as a weight on the species of structure $\tilde{\mathcal{G}}$.

In order to interpret the species $\tilde{\mathcal{G}}$ in more detail, we look at how it is built up. We can imagine partitioning the collection of objects in $\tilde{\mathcal{G}}$ with representatives of the form $(\emptyset, S)$. The representative $(\emptyset, S)$ corresponds to all the elements of $\tilde{\mathcal{G}}$ which are written, for any subset $K \subset S$, $(K, S \setminus K)$. This means that we have the identity:

$$\tilde{\mathcal{G}} = \mathcal{P} \circ \mathcal{E}^{\bullet} \tag{VI.7.18}$$

where $\mathcal{P}$ is the species corresponding to subsets and $\mathcal{E}$ is the set species.

The interpretations and inductive approach becomes most useful when we are dealing with edge-multiplicative weights. This is the key idea in the bounds being effective and having the corresponding identity. Edge multiplicativity means that we have the identity:

$$w(D_{\{i\}}\mathcal{G}[J]) = \left( \sum_{K \subset J} \prod_{j \in K} \zeta_{i,j} \right) w(\mathcal{G}[J]) \tag{VI.7.19}$$

This motivates us to make the definition of $w_i(S)$ for $S \in \mathcal{P}[J]$ as $\prod_{j \in S} \zeta_{i,j}$, so that we may write the edge-multiplicativity rule as:

$$w(D_{\{i\}}\mathcal{G}[J]) = w_i(\mathcal{P}[J])w(\mathcal{G}[J]) \tag{VI.7.20}$$

We may also write for the inductive relationship (VI.7.17):

$$w(\mathcal{G}^{\star(-1)} \star D_I\mathcal{G}[J]) = \sum_{K \subset J} w(\mathcal{G}^{\star(-1)}[J \setminus K])w_i(\mathcal{P}[I' \cup K])w(\mathcal{G}[I' \cup K]) \tag{VI.7.21}$$

We may also make a connection here with the Kirkwood Salsburg equations, which in this notation may be written as:

$$\tilde{\phi}_I[J] = \exp(-\beta W^{\tau(I)}(I')) \sum_{S \subset J} K(I', S)\tilde{\phi}_{I' \cup S}(Y \setminus S) \tag{VI.7.22}$$

This prefactor is precisely $w_{\tau(I)}(\mathcal{P}[I'])$. We wish now to understand the Kirkwood-Salsburg kernel as a species of structure. Recall that the Kirkwood Salsburg kernel

is defined by:

$$K((x)_m,(y)_n) = \prod_{j=1}^{n}\left(\prod_{i=1}^{m}\exp(-\beta\Phi(y_i-x_j))-1\right)$$
$$= \prod_{j=1}^{n}\left(\prod_{i=1}^{m}(1+\zeta_{i,j})-1\right) \tag{VI.7.23}$$

This can be interpreted as the collection of graphs where for each element of $(y)_n$ we have *at least* one element of $(x)_m$ connecting to it. This corresponds to the non-empty subset species $\mathcal{P}_+[(x)_m]$ for each $y_j$.

## Conclusions & Open Questions

The theory of combinatorial species of structure allows one to simply understand the whole structure of many of the elements of statistical mechanics and in particular the virial and cluster expansions. The ability to cast the relationships in the form given by the combinatorial species of structure, allows for an ease of understanding where the coefficients come from without the need for arduous derivations related to looking for simple connections between different coefficients. The techniques used in this chapter inform further uses of the disymmetry theorem and Lagrange inversion in the multispecies case in Chapter IX. Furthermore, the connections between the two functions and the interpretations as tree fixed point equations are indicated in Chapter VII.

Questions that are still left open are the degree to which combinatorial connections can be made and how these can assist with providing bounds for expansions or whether it provides insights to related expansions in mathematical physics in general.

# Chapter VII

# Graph-Tree Identities and Inequalities

The original contribution made by this chapter is to indicate the interpretations of the graph-tree identities in light of Chapter III and to provide interpretations of the literature presented in this chapter for two-connected graphs. Furthermore, in Section VII.2, the focus is on using these interpretations to gain an identity, interpretation and bound for the case of two-connected graphs.

Graph-tree inequalities and identities and their application to statistical mechanics originate from the work of Brydges Battle and Federbush [Bat84, BaFe84, BrFe78]. In Section I.2, through Mayer's First Theorem, we have that the cluster expansion can be interpreted as the exponential generating series of weighted connected graphs:

$$\beta P = \sum_{n \geq 1} \frac{z^n}{n!} \sum_{g \in \mathcal{C}[n]} w(g) \tag{VII.0.1}$$

A naïve approach for understanding the convergence of this series is to make the uniform upper bound for $g \in \mathcal{C}[n]$:

$$|w(g)| \leq \sup_{h \in \mathcal{C}[n]} |w(h)| = A_n \tag{VII.0.2}$$

This leads to considering convergence of the series:

$$\sum_{n \geq 1} \frac{|z|^n}{n!} l_n A_n \tag{VII.0.3}$$

where $l_n$ is the number of connected graphs on $n$ vertices. From Flajolet and Sedgewick [FlSe09], we see that $\frac{l_n}{2^{n^2/2}} \to 1$. This can be seen from the fact that

we can provide a lower bound on the number of connected graphs, by considering starting from a tree and then realising that we may add or not any missing edge from this tree. This gives a lower bound of $2^{\binom{n}{2}-(n-1)} = 2^{\binom{n-1}{2}}$. We have the upper bound from the total number of graphs as $2^{\binom{n}{2}}$. Hence the ratio tends to 1. We note that $A_n$ is at best geometric in $n$ and so we will have zero radius of convergence through this approach.

Furthermore, for two-connected graphs, we realise we may obtain a similar lower bound, by realising that we can start with an $n$-cycle on all the vertices and add or not any of the remaining edges, giving a lower bound of $2^{\binom{n}{2}-n} = 2^{\binom{n-1}{2}-1}$. Hence we obtain the same ratio in the limit and have the same problems as for connected graphs.

However, there are some important observations to make about the weight functions for connected graphs. The signs of the weights $w(g)$ (for positive potentials) alternate in the number of edges of the graph $g$. We therefore need to understand how cancellations may occur due to the alternating signs.

The key approach for connected graphs is to find a majorising tree expansion. In particular for the Tonks gas and the one particle hard core gas, we have seen in Chapter III, through the work of Bernardi [Ber08], that we precisely obtain contributions from certain classes of trees.

Graph-tree identities [AbRi94, Bat84, BrFe78, Bry84, Pen67] provide the identity:

$$\sum_{g \in \mathcal{C}[n]} w(g) = \sum_{t \in \mathfrak{a}[n]} \tilde{w}(t) \tag{VII.0.4}$$

where $\tilde{w}$ are appropriately modified weights.

We then may proceed by using the triangle inequality:

$$\left| \sum_{t \in \mathfrak{a}[n]} \tilde{w}(t) \right| \leq t_n U_n \tag{VII.0.5}$$

where $U_n$ is a uniform bound on $|\tilde{w}(g)|$ for $g \in \mathfrak{a}[n]$. This form of rewriting series of connected graphs in terms of trees is also useful in the subject of renormalisation in quantum field theory. It is most notably a key part of loop vertex expansions [RiWa10a, RiWa10b].

The idea of Brydges and Federbush's tree graph identity [BrFe78] is to repeatedly use the fundamental theorem of calculus with an interpolating version of the potential in the exponential: $\exp(-\beta \mathcal{U})$. The interpolating parameters $(s_i)$ introduced play the rôle of separating the first $i$ particles from the rest, thereby

splitting the exponential according to the interactions. This approach is considered in Section VII.1 and is hoped can be reproduced for two-connected graph functions. However, the terms become a lot more complicated in this case.

Further improvements of the exposition of graph-tree identities and the estimates were also made by Battle and Federbush [Bat84, BaFe84]. In the paper by Abdesselam and Rivasseau [AbRi94], a further more symmetric identity is produced, which relies on combinatorial arguments and not the inductive approach from the method of Brydges and Federbush. The symmetry of the identity provides a conceptual aesthetic to this theory, making the combinatorics clear and explicit. There are also connections of these expansions to Kruskal's algorithm, where one desires to find a graph tree identity respecting the stability condition of the potential. These are found in the notes of Helmuth [Hel14] and the paper of Rivasseau and Tanasa [RiTa14]. A key point is that there is a connection between the Penrose construction and these graph-tree identities.

In Section VII.3, the construction of Penrose [Pen67] is presented. The presentation of this construction is taken from the paper of Fernández and Procacci [FePr07]. It is connected to the matroid concepts presented later in the chapter. In this case, the possibility of many different related constructions is conveyed and the modified weights are obtained. Of further importance are the connections to Bernardi's involutions [Ber08], given in Chapter III. This approach is particularly useful for positive potentials and has been generalised to stable potentials. The paper of Poghosyan and Ueltschi [PoUe09] makes use of this and the paper by Procacci [Pro07] in order to obtain cluster expansion bounds in various contexts. This is the motivation for Chapter VIII and is used in connection with expressions of the virial coefficients in terms of the cluster coefficients to obtain a tree bound for the virial coefficients.

The cluster expansion bounds of Fernández and Procacci in Theorem I.5.3, can be understood in terms of tree operators. This has been refined and generalised in the work of Temmel [Tem12]. The key connection here is the fact that the alternative methods of achieving cluster expansion bounds through fixed point equations also rely on tree type estimations.

The notion of a tree-graph identity has also been generalised to that of an analogous identity for matroids in the paper by Faris [Far12a]. There are also key connections to the $q$-state Potts model [Sok05]. The key ideas of the approaches to the tree graph identities generalise nicely in the framework of matroids and can be understood on this level. The advantage of this more general framework is that it provides useful connections to the proofs of Chapter III. The combinatorial

understanding of the graphs in the form of matroids is closely linked to methods of achieving bounds on cluster coefficients and provides insight on an adaptation to understand bounds on virial coefficients. This generalisation is presented in Section VII.4. Matroids however are not the right structure to use for two-connected graphs since they require a basis, where all the sets are the same size, but minimal two-connected graphs all have different numbers of edges. The target is to have a partially ordered set with some more structure that is sufficient to have the concepts of internally and externally active edges as in the work of Sokal [Sok05].

## VII.1 Tree Graph Identities and the Fundamental Theorem of Calculus

### VII.1.1 The Brydges-Federbush Formula

The tree-graph identities arising from the paper of [BrFe78] have connections to more general contexts in which one can create a cluster expansion. The ideas are present in the paper on the Hamilton-Jacobi Equation [BrKe87] and a version of the identity is used in Quantum lattice systems [KoRe96], phase cell cluster expansions [BaFe82, BaFe83] and loop vertex expansions of renormalisation [RiWa10a, RiWa10b].

This is also applied in the paper by Brydges and Imbrie [BrIm03], where tree identities are used to perform dimensional reduction from a branched polymer model to a hard core gas, with the dimension reduced by two. This comes from recognising the trees in the cluster expansion of the logarithm of the partition function for the hardcore gas and realising that branched polymers correspond to such trees.

This section indicates this formula and how it is proved and emphasises key links to the rest of the thesis.

The notation $\Phi_{i,j} := \Phi(x_i - x_j)$ is used to denote the pair potential between particles $i$ and $j$.

The application of the fundamental theorem of calculus used in [BrFe78] is:

$$\exp(f(1)) = \exp(f(0)) + \int_0^1 \mathrm{d}s\, f'(s) \exp(f(s)) \tag{VII.1.1}$$

This is applied repeatedly in order to obtain tree estimates for connected graph sums.

We recall the Mayer series (Section I.2) for pressure in terms of activity:

$$\beta P = \sum_{n=1}^{\infty} b_n z^n \tag{VII.1.2}$$

In this section, for a translation and rotation invariant potential $\Psi$, we make the following assumptions:

**Assumption 8** (Stability)**.** *That* $\exists\, B \geq 0$ *such that* $\forall n$ *and* $\forall (x_1, \cdots, x_n) \in (\mathbb{R}^D)^n$*, we have:*

$$\sum_{1 \leq i < j \leq n} \Phi(x_i - x_j) \geq -Bn \tag{VII.1.3}$$

**Assumption 9** (Integrability)**.**

$$\|\Phi\|_1 = \int |\Phi(x)| \, \mathrm{d}^3 x < \infty \tag{VII.1.4}$$

**Remark 38** (Temperedness)**.** *This is a modified version of temperedness found in Section II.3 and is precisely the form used by Procacci [Pro07] and Poghosyan and Ueltschi [PoUe09], who use this version of the graph-tree identity. This formulation of temperedness arises from the fact that the potential is brought outside of the exponential through the repeated use of the fundamental theorem of calculus.*

The technique is straightforward and relies on interpolating or separating the particles with indeterminates $s_i$ for $i \in [n]$. The $j$th indeterminate, $s_j$, separates the first $j$ particles from the remaining particles. The technique focuses on the partition function integrals $Z_\Lambda^{(n)}$ for $n$ particles. Once we have the technique for these functions it generalises nicely to the connected functions. We start with the three-particle partition function:

$$Z_\Lambda^{(3)} = \frac{1}{3!} \int_{\Lambda^3} \exp(-\beta(\Phi_{1,2} + \Phi_{1,3} + \Phi_{2,3})) \, \mathrm{d}x_1 \, \mathrm{d}x_2 \, \mathrm{d}x_3 \tag{VII.1.5}$$

The function $f(1)$ in this example is $-\beta(\Phi_{1,2} + \Phi_{1,3} + \Phi_{2,3})$. We want to separate the first particle from the rest, so our interpolating indeterminate $s_1$ is multiplied by the terms containing 1. That is:

$$f(s_1) = -\beta(s_1(\Phi_{1,2} + \Phi_{1,3}) + \Phi_{2,3}) \tag{VII.1.6}$$

This gives:

$$\exp(-\beta(\Phi_{1,2}+\Phi_{1,3}+\Phi_{2,3})) = \exp(-\beta\Phi_{2,3}) + \int_0^1 \mathrm{d}s_1\,(-\beta)(\Phi_{1,2}+\Phi_{1,3}) \times \exp(-\beta(s_1(\Phi_{1,2}+\Phi_{1,3})+\Phi_{2,3})) \tag{VII.1.7}$$

These functions are integrated over the whole space and so interchanging the labels 2 and 3 in the second term, we obtain:

$$2\int_0^1 \mathrm{d}s_1(-\beta)\Phi_{1,2}\exp(-\beta(s_1(\Phi_{1,2}+\Phi_{1,3})+\Phi_{2,3})) \tag{VII.1.8}$$

We now use $s_2$ to separate the third particle from the first two, so that equation (VII.1.8) becomes:

$$2\int_0^1 \mathrm{d}s_1\,(-\beta)\Phi_{1,2}\exp(-\beta s_1\Phi_{1,2}) + 2\int_0^1 \mathrm{d}s_1 \int_0^1 \mathrm{d}s_2\,(-\beta)^2\Phi_{1,2} \times (\Phi_{2,3}+s_1\Phi_{1,3})\exp(-\beta W(s_1,s_2)) \tag{VII.1.9}$$

where,

$$W(s_1,s_2) = (s_1\Phi_{1,2} + s_1s_2\Phi_{1,3} + s_2\Phi_{2,3}) \tag{VII.1.10}$$

For the first term in (VII.1.7), we perform the same interpolation to obtain:

$$1 + (-\beta)\int_0^1 \mathrm{d}s_2\,\Phi_{2,3}\exp(s_2\Phi_{2,3}) \tag{VII.1.11}$$

We replace $s_2$ with $s_1$ and $\Phi_{2,3}$ by $\Phi_{1,2}$, to obtain the symmetric equation:

$$Z_\Lambda^{(3)} = \frac{1}{3!}(1 + 3\int_0^1 \mathrm{d}s_1\,\Phi_{1,2}\exp(-\beta s_1\Phi_{1,2}) + 2\int_0^1 \mathrm{d}s_1\int_0^1 \mathrm{d}s_2\,(-\beta)^2\Phi_{1,2}(s_1\Phi_{1,3}+\Phi_{2,3})\exp(-\beta W(s_1,s_2)) \tag{VII.1.12}$$

We define the connected function $b_n^\Lambda$ as the sum over weighted connected graphs:

$$K_n^\Lambda = \frac{1}{n!}\int_{\Lambda^n} \sum_{g\in\mathcal{C}[n]} \prod_{\{i,j\}\in E(g)} (\exp(-\beta\Phi_{i,j})-1)\,\mathrm{d}x_1\cdots\mathrm{d}x_n \tag{VII.1.13}$$

In order to generalise this, the interpolating function $W$ inside of the exponential in each term needs to be defined. To the set $[l]$ of $l$ particles, $l>1$, we associate $\sigma_{l-1}$ a sequence of parameters $(s_1,\cdots s_{l-1})$ and a potential, $W^{[l]}(\sigma_{l-1})$, defined by an inductive process:

$$W_0^{[l]} = V^{[l]} \tag{VII.1.14}$$

$$W_i^{[l]} = (1-s_i)W_{i-1}^{[l],[i]} + s_i W_{i-1}^{[l]}, i=1,\cdots,l-1 \tag{VII.1.15}$$

$$W_{l-1}^{[l]} = W^{[l]}(\sigma_{l-1}) \tag{VII.1.16}$$

Where

$$V^{[l]} = \sum_{y\in[l]^{(2)}} \Phi_y \text{ and } W^{X,X'} = W^{X'} + W^{X\setminus X'}$$

We start with the unaffected potential (VII.1.14) and then at step $i$, separate the pair potentials according to whether they connected the first $i$ particles to the rest or not. If they do they are multiplied by $s_i$ as is indicated in (VII.1.15). We must repeat this $l-1$ times where we obtain the final expression (VII.1.16).

We realise that:

$$W^{[l]}(\sigma_{l-1}) = \sum_{1\le i<j\le l} s_i\cdot s_{i+1}\cdots s_{j-1}\Phi_{i,j} \tag{VII.1.17}$$

In the process of separating the particles, we have a choice upon separating the $i$th particle, which of the terms $\Phi_{j,i}$ for $j<i$, we bring down from the exponential in each term. We represent this choice by a function $\eta:[2,l]\to[l-1]$, satisfying $\eta(i)<i$. The term $\Phi(\eta(i),i)$ appears in the term corresponding to the given $\eta$. We represent this function as a tree graph, where $i$ is connected to $\eta(i)$, as is displayed in Figure VII.1. We certainly have a tree, since we have $l-1$ edges and the graph is connected. From each vertex $i>1$, there is an edge to a strictly smaller vertex. We may follow such edges down until we get to 1. Every vertex is therefore connected to 1 so the graph is connected. All connected graphs on $[l]$ with $l-1$ edges are tree graphs.

Define the measure:

$$\int \mathrm{d}\sigma_{l-1} = \int_0^1 \mathrm{d}s_1 \cdots \int_0^1 \mathrm{d}s_{l-1} \tag{VII.1.18}$$

Introduce the tree function $f$ depending on the graph $\eta$ and the interpolation variables $\sigma_{l-1}$, by:

$$\begin{aligned} f(\eta, \sigma_{l-1}) &= \prod_{i=2}^{l} s_{i-1} s_{i-2} \cdots s_{\eta(i)} \text{ for } l > 2 \\ f(\eta, \sigma_1) &= 1 \end{aligned} \tag{VII.1.19}$$

The empty product is understood to be 1. The connected version, $b_n^\Lambda$, is written as:

$$b_n^\Lambda = \sum_\eta \xi_n^\Lambda(\eta) \tag{VII.1.20}$$

The sum is over all $\eta$ satisfying $\eta(i) < i$, which is viewed as a subset of trees. In particular it is the collection of rooted increasing trees. The expression for a given $\eta$ is:

$$\xi_n^\Lambda(\eta) = \frac{(-\beta)^{n-1}}{l} \int \mathrm{d}\sigma_{n-1} \int_\Lambda \mathrm{d}x^{(n)}\, f(\eta, \sigma_{n-1}) \exp(-\beta W^{(n)}(\sigma_{n-1})) \prod_{i=2}^{n} \Phi_{i,\eta(i)} \tag{VII.1.21}$$

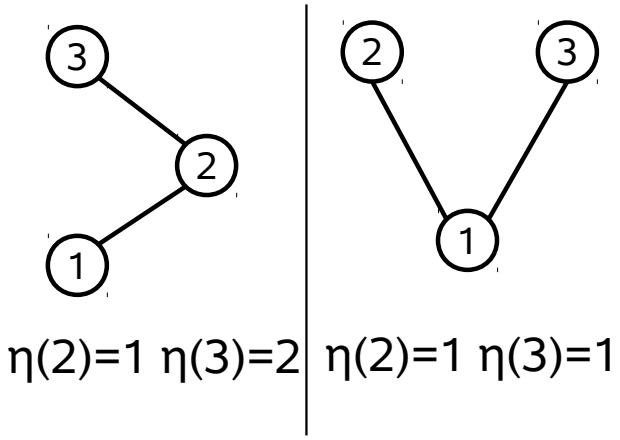



Figure VII.1: Two Trees Corresponding to the Different $\eta$ Functions for $l = 3$

To achieve the bounds in determining the radius of convergence of the cluster

expansion, we need to bound the individual terms $\xi_n^\Lambda(\eta)$

**Proposition VII.1.1** (Brydges-Federbush)**.** *We have the following estimate on the integral over the s-parameters in (VII.1.21):*

$$\sum_\eta \int \mathrm{d}\sigma_{n-1}\, f(\eta, \sigma_{n-1}) \leq \exp(n-1) \tag{VII.1.22}$$

*Proof.* This is shown by first noting that the left hand side of (VII.1.22) is less than or equal to:

$$\int_0^1 \mathrm{d}s_1 \cdots \int_0^1 \mathrm{d}s_{n-1} \sum_\eta f(\eta, \sigma_{n-1}) \exp\left(\sum_{i=1}^{n-1} s_{n-1}s_{n-2}\cdots s_i\right) \tag{VII.1.23}$$

In order to see the $\exp(n-1)$ factor, consider first the largest label $n$ in the increasing tree $\eta$, which is necessarily a leaf. $s_{n-1}$ appears as an overall factor in the exponent and in each tree. Each of the accompanying factors in the exponent indicates the different ways in which this leaf can be added to the tree without $n$, with the smallest subscript of $s$ appearing in the individual product indicating the vertex it is attached to and thus it appears as a factor in the sum over increasing trees. When the integral is performed over $s_{n-1}$, this factor corresponding to adding the vertex $n$ is removed from the integrand and we obtain:

$$\int_0^1 \mathrm{d}s_1 \cdots \int_0^1 \mathrm{d}s_{n-2} \sum_\eta f(\eta, \sigma_{n-2}) \exp\left(1 + \sum_{i=1}^{n-2} s_{n-2}\cdots s_i\right) \tag{VII.1.24}$$

where the sum is now over increasing trees on $n-1$ vertices. We obtain a factor of $\exp(1)$ from each successive integral. □

**Proposition VII.1.2** (Brydges-Federbush)**.** *The total sum over the tree contributions gives the estimate for the connected functions as:*

$$|c_n^\Lambda| \leq \beta^{n-1}(\|\Phi\|_1)^{n-1} \exp(n\beta B)\frac{\exp(n-1)}{n}|\Lambda| \tag{VII.1.25}$$

*Proof.* We have the stability bound (VII.1.3) and the fact that $W^{(n)}(\sigma_{n-1})$ is a convex combination of the individual pair potential summands, where at each stage in (VII.1.15), we see that the first term is bounded below by $(1-s_i)\times -B(n-i) - Bi = -Bn(1-s_i)$ and the second term is inductively bounded below by $-s_iBn$ so that we have at each stage $W_i^{[n]} \geq -Bn$, which gives a uniform upper bound $\exp(\beta Bn)$

in (VII.1.21). We are now left to bound the integral:

$$\int_{\Lambda^l} \mathrm{d}x^{(n)} \prod_{i=2}^{n} \left|\Phi_{i,\eta(i)}\right| \tag{VII.1.26}$$

We may make the change of variables $w_i = x_i - x_{\eta(i)}$ for $2 \leq i \leq n$ and $w_1 = x_1$. We therefore achieve the upper bound of $\|v\|_1^{n-1}|\Lambda|$ for the integrals. Putting these together gives the bound (VII.1.25). □

This bound on the cluster coefficients, leads to the bound on the radius of convergence of the cluster expansion $\mathcal{R}_{\mathrm{May}}$ as:

$$\mathcal{R}_{\mathrm{May}} \geq \frac{1}{\beta e \|\Phi\|_1} \exp(-\beta B) \tag{VII.1.27}$$

This different bound is used in the paper [Tate13] to formulate different bounds for the virial expansion, as is indicated in Section IV.6.

It is useful to emphasise the connection here to the paper of Bernardi [Ber08]. The edges are given a partial order in the case of this fundamental theorem of calculus approach. We see that since $s_i \in [0,1]$ that the prefactor of an edge $\Phi_{i,j}$ has as many $s$-factors as the distance $|i-j|$. If we write edges uniquely as $(a,b)$ with $a < b$, we see that this corresponds to a partial order: $(i,j) \preccurlyeq (k,l)$ if and only if $i \leq k$ and $l \leq j$. That is edges are ordered according to the size of their prefactors when they can be easily compared. What is useful to understand further is how this partial order on edges can help create fundamental theorem of calculus type results for the partially ordered set of connected graphs.

### VII.1.2 Combinatorial Graph-Tree Expansion

The fundamental theorem of calculus identities given by Brydges and Federbush are further generalised into a neat symmetric formula by Abdesselam and Rivasseau in [AbRi94].

**Definition** (Unordered Forests and Trees)**.** *A graph $\mathfrak{F} = \{l_1, \cdots, l_\tau\}$ containing no loops is called a u-forest, or unordered forest. A tree $\mathfrak{T}$ is a connected u-forest.*

**Definition** (Clusters)**.** *The supports of the disconnected trees making up a u-forest $\mathfrak{F}$ are called connected components or clusters. These include isolated points.*

**Theorem VII.1.3** (Brydges-Kennedy Graph-Tree Identity)**.** *For indeterminates $(u_l)_{l \in [n]^{(2)}}$, labelled by edges in $[n]^{(2)}$, which can be substituted for by elements of a*

*commutative Banach algebra $\mathcal{B}$, we have the following identity:*

$$\exp\left(\sum_{l\in[n]^{(2)}} u_l\right) = \sum_{\substack{\mathfrak{F}=\{l_1,\cdots,l_\tau\}\\ u\text{-}forest}} \left(\prod_{\nu=1}^{\tau}\int_0^1 \mathrm{d}h_{l_\nu}\right)\left(\prod_{\nu=1}^{\tau} u_{l_\nu}\right)\exp\left(\sum_{l\in[n]^{(2)}} h_l^{\mathfrak{F}}(\mathbf{h})u_l\right) \tag{VII.1.28}$$

*where we sum over all possible values of $\tau$, including $\tau = 0$ corresponding to the empty forest and contributing* 1. *The function $h^{\mathfrak{F}}_{\{ij\}}(\mathbf{h}) = \inf\{h_l | l \in L_{\mathfrak{F}}\{ij\}\}$, where $L_{\mathfrak{F}}\{ij\}$ is the unique path in the forest $\mathfrak{F}$ connecting $i$ to $j$. If no such path exists, then $h^{\mathfrak{F}}_{\{ij\}}(\mathbf{h}) = 0$.*

To understand how it relates to graphs, it is instructive to cast this in the form corresponding to Mayer's approach of replacing $\exp(u_l)$ with $(1 + f_l)$. This transforms the LHS of (VII.1.28) into $\sum_{E\subset[n]^{(2)}}\prod_{l\in E} f_l$, which we can cast in the familiar form: $\sum_{g\in\mathcal{G}[n]}\prod_{\{ij\}\in E(g)} f_{\{ij\}}$.
When we use Möbius inversion on this expression, we find that the corresponding identity for the sum over connected graphs involves trees instead of u-forests. This formula captures a symmetric version of using Penrose's construction.

## VII.2 Applying Bounds to the Two-Connected Case

In order to create a tree expansion for two-connected graphs, one method is to look at how the virial coefficients can be written in terms of cluster coefficients through Lagrange inversion. The key idea will then be to see how substituting the tree expansion for the cluster coefficients leads to a modified tree expansion for virial coefficients. The key idea is how the combinatorial factors for the particular terms can be interpreted.

The easy way is to to see cluster coefficients written in terms of the virial coefficients. This was first understood in the work of Mayer [MMay40]. A key idea in obtaining the expression of connected graphs in terms of two-connected graphs was the concept of a Husimi graph. It is taken for a definition that Husimi graphs are precisely those connected graphs whose blocks are all complete graphs. The number of such graphs was understood in the papers of Ford and Uhlenbeck [FoUh56a, FNU56, FoUh56b], where the graphs were categorised by the number of blocks of a certain size that were present. This was also emphasised in the paper of Leroux [Ler04]:

$$b_n = \sum_{k=1}^{n} \sum_{\substack{\{n_i\}_{i=2}^{n} \\ \sum_{i=2}^{k}(i-1)n_i = n-1}} \mathrm{Hu}(n_2, n_3, \cdots) \prod_{i=2}^{\infty} (\beta_{i-1})^{n_i} \tag{VII.2.1}$$

Where:

$$\mathrm{Hu}(n_2, n_3, \cdots) = \frac{(n-1)! n^{\sum n_j - 1}}{\prod_{i \geq 2} (i-1)!^{n_i} n_i!} \tag{VII.2.2}$$

These relations could be inverted by hand, but an easier method is to use Lagrange inversion through a Cauchy integral representation, as was used in Chapter IV. This is given by:

$$c_n = \frac{\beta}{2n\pi i} \oint \frac{\mathrm{d}z}{z\rho^{n-1}} \tag{VII.2.3}$$

We can consider the fugacity expansion for density and write it as:

$$\rho = z(1 + A(z)) \tag{VII.2.4}$$

where:

$$A(z) = \sum_{n=2}^{\infty} n b_n z^{n-1} \tag{VII.2.5}$$

If we substitute this into (VII.2.3), we get:

$$\begin{aligned} c_n &= \frac{1}{2n\pi i} \oint \frac{\mathrm{d}z}{z^n (1 + A(z))^{n-1}} \\ &= \frac{1}{2n\pi i} \oint \frac{\mathrm{d}z}{z^n} \sum_{k=0}^{\infty} \binom{n+k-2}{k} (-1)^k A(z)^k \end{aligned} \tag{VII.2.6}$$

The equation (VII.2.6) can be expressed as identifying the $z^{n-1}$ coefficient of the expansion in terms of $A(z)$. We note that since $A(z)$ has lowest term $z$, we only need to consider terms up to $k = n-1$ and we can omit the 0 from the sum as this is a constant term.

We thus obtain:

$$\begin{aligned} c_n &= [z^{n-1}] \frac{1}{n} \sum_{k=1}^{n-1} \binom{n+k-2}{k} (-1)^k A(z)^k \\ &= \frac{1}{n} \sum_{k=1}^{n-1} \binom{n+k-2}{k} (-1)^k \sum_{(l_i)_{i=1}^{k}} \prod_{i=1}^{k} (l_i + 1) b_{l_i+1} \end{aligned} \tag{VII.2.7}$$

where the sum is over $(l_i)_{i=1}^k$ satisfying $\sum_{i=1}^{k} l_i = n-1$.

We can recast the final sum in (VII.2.7) in terms of $m_i :=$ number of $l_j+1 = i$. The number of such sequences with a given collection of $\{m_i\}$ is $\binom{k}{m_2 \cdots m_{n-1}}$. Thus we get the expression:

$$c_n = \frac{1}{n} \sum_{k=1}^{n-1} \binom{n+k-2}{k} (-1)^k k! \sum_{(m_i)_{i=1}^{n-1}} \prod_{i=2}^{n-1} \frac{(ib_i)^{m_i}}{m_i!} \tag{VII.2.8}$$

where the sum is over sequences $(m_i)_{i=1}^{n-1}$, satisfying:

$$\sum_{i=2}^{n-1} (i-1) m_i = n-1 \text{ and } \sum_{i=1}^{n-1} m_i = k \tag{VII.2.9}$$

Recall $\beta_{n-1} = -\frac{n-1}{n!} c_n$ and so we achieve the relationship for the two-connected graphs $\beta_n$:

$$\beta_{n-1} = -\frac{n-1}{n} \sum_{k=1}^{n} \binom{n+k-1}{k} (-1)^k \frac{k!}{n!} \sum_{(m_i)_{i=2}^{n}} \prod_{i=2}^{n} \frac{(ib_i)^{m_i}}{m_i!} \tag{VII.2.10}$$

We may then write each $b_n$ in terms of its tree version. If we do this, we have to then understand how the combinatorics can give a useful labelling of the trees in the products and how they relate to the full trees. The key problem is that the exponential and the $f(\eta, \sigma_{n-1})$ functions will not match up and there won't be any straightforward cancellations. Instead we will achieve a usual tree expansions but with some awkward factors.

Given a single term in the expansion (VII.2.10), we see it is formed from a product of $b_i$. From each $b_i$ factor we obtain a possible increasing tree. We want to see how to understand the product of these increasing trees on certain label sets in terms of an increasing tree on a different label set. To this end we understand the distribution of the labels amongst the trees in the product. Given a collection of these trees, we determine the distribution of labels amongst the vertices, firstly, by ordering the connected graph components. This deals with the $\frac{k!}{\prod_{i=2}^{n} m_i!}$ factor. If we write the ordered connected graph sizes as $i_1, i_2, \cdots i_k$, then the labels $1, \cdots i_1$ are given to the first tree with the same order as the tree has. The second tree is given the labels $i_1, \cdots, i_1 + i_2 - 1$. The picture is that each subsequent increasing tree is given the following labels so that its root is the largest label of the previous tree.

We can therefore make an identification in each term with an increasing

tree. The $f$-factors depend only upon the progeny of each vertex,which is included entirely within one factor. Therefore, the $f$-factor will be the same for all of these trees. The factors in the exponential, will differ depending on how we have split the graph. This still needs to be understood more fully or whether it can be cast in a nice way through some of the more symmetric procedures explained later.

### VII.2.1 Alternative Graph Tree Identities and the Hardcore Case

The alternative presentations of the tree graph identity are particularly useful in certain circumstances. These can be found in [Pro07] and the key connection is with the presentation of the matroid version of these identities in Section VII.4.

**Lemma VII.2.1** (Probability Measure of Tree Integrals- [Pro07, BaFe84])**.** *The integral over trees forms a probability measure, that is, for all tree graphs $\tau$:*

$$\prod_{j=1}^{n-1}\left(\int_0^1 \mathrm{d}t_j\right)\sum_{(X_1\cdots X_{n-1})_\tau} t_1^{b_1-1}\cdots t_{n-1}^{b_{n-1}-1}=1 \tag{VII.2.11}$$

**Theorem VII.2.2** (Alternative Presentation of the Tree Graph Identity)**.** *The tree graph identity may be written, for a potential $\Phi_{i,j}$ as:*

$$\begin{aligned}\sum_{g\in\mathcal{C}[N]}\prod_{\{i,j\}\in E(g)}\left(e^{-\Phi_{i,j}}-1\right)=\sum_{\tau\in\mathfrak{a}[N]}\prod_{\{i,j\}\in E(\tau)}(-\Phi_{i,j})\left(\prod_{j=1}^{n-1}\left(\int_0^1\mathrm{d}t_j\right)\right)\\ \sum_{(X_1,\cdots,X_{n-1})_\tau}\prod_{j=1}^{n-1}t_j^{b_j-1}\exp(W((X)_{[n-1]},t_{[n-1]})\end{aligned} \tag{VII.2.12}$$

$X_{[n-1]}=(X_1,\cdots,X_{n-1})$ is a sequence of increasing subsets of $[N]$, such that $X_1=\{1\}$ and $X_i\subset X_{i+1}$ with $|x_i|=i$. A sequence $(X_1,\cdots X_{n-1})$ is compatible with $\tau$, written $(X_1,\cdots X_{n-1})_\tau$, if for all $i$, $X_i$ contains $i-1$ edges of $\tau$. The edge $\{j,k\}$ is said to cross $X_i$ if $k\in X_i$ but $j\notin X_i$ for some $i$. $b_i=$ the number of edges of $\tau$, which cross $X_i$ and the $W$-function is defined as:

$$W((X)_{[n-1]},(t)_{[n-1]})=\sum_{1\le j<k\le n}t_1(\{j,k\})\cdots t_{n-1}(\{j,k\})\Phi_{j,k} \tag{VII.2.13}$$

where

$$t_i(\{j,k\})=\begin{cases}t_i\in[0,1] & \text{if } \{j.k\} \text{ crosses } X_i\\ 1 & \text{otherwise}\end{cases} \tag{VII.2.14}$$

Generalised stability is defined by:

**Definition** (Generalised Stability)**.** *The collection* $(\Phi_{i,j})_{1\leq i<j\leq N}$ *satisfies the generalised stability property, if for any subset* $S \subset [N]$*, we have constants* $(\Phi_{i,i})_{i\in[N]}$*, which satisfy:*

$$\sum_{i\in S} \Phi_{i,i} + \sum_{\{i,j\}\in S^{(2)}} \Phi_{i,j} \geq 0 \tag{VII.2.15}$$

Generalised stability then implies that the interpolating function satisfies:

$$W((X)_{[n-1]}, t_{[n-1]}) \geq -\sum_{i=1}^{n} \Phi_{i,i} \tag{VII.2.16}$$

If we have a uniform upper bound for the $V_{i,i}$, say $B$, then we have the usual form of stability and obtain the following inequalities:

**Lemma VII.2.3** (Tree-Graph Inequalities)**.** *For a stable potential* $\Phi$*, with no hard core interactions, we have the following inequality:*

$$\left| \sum_{g\in\mathcal{C}[N]} \prod_{\{i,j\}\in E(g)} \left(e^{-\Phi_{i,j}} - 1\right) \right| \leq e^{nB} \sum_{\tau\in\mathfrak{a}[N]} \prod_{\{i,j\}\in E(\tau)} |\Phi_{i,j}| \tag{VII.2.17}$$

*For repulsive potentials with a hardcore, we may write the inequality as:*

$$\left| \sum_{g\in\mathcal{C}[N]} \prod_{\{i,j\}\in E(g)} \left(e^{-\Phi_{i,j}} - 1\right) \right| \leq \sum_{\tau\in\mathfrak{a}[N]} \prod_{\{i,j\}\in E(\tau)} \left|e^{-\Phi_{i,j}} - 1\right| \tag{VII.2.18}$$

The modification described by Procacci [Pro07] is for polymer systems but it is possible to make the transition to continuous systems [PoUe09]. We emphasise that this is for a finite polymer set $\mathfrak{F}$.

The interacting partition function is:

$$\Xi_{\mathfrak{F}}(\mathbf{z}) = 1 + \sum_{n\geq 1} \frac{1}{n!} \sum_{(\gamma_1,\cdots,\gamma_n)\in\mathfrak{F}^n} z_{\gamma_1}\cdots z_{\gamma_n} e^{-\sum\limits_{1\leq i<j\leq n} V(\gamma_i,\gamma_j)} \tag{VII.2.19}$$

From which we achieve the standard cluster expansion:

$$\log \Xi_{\mathfrak{F}}(\mathbf{z}) = \sum_{n=1}^{\infty} \frac{1}{n!} \sum_{(\gamma_1,\cdots,\gamma_n)\in\mathfrak{F}^n} \phi^T(\gamma_1,\cdots,\gamma_n) z_{\gamma_1}\cdots z_{\gamma_n} \tag{VII.2.20}$$

where:

$$\phi^T(\gamma_1, \cdots, \gamma_n) := \begin{cases} 1 & \text{if} n = 1 \\ \sum_{g \in \mathcal{C}[n]} \prod_{\{i,j\} \in E(g)} \left(e^{-V(\gamma_i, \gamma_j)} - 1\right) & \text{for } n \geq 2 \end{cases} \tag{VII.2.21}$$

In order for the partition function to be analytic, we require stability:

**Definition** (Polymer Stability)**.** *A polymer interaction $V$ is called stable if for any fixed $n \in \mathbb{N}$ and $(\gamma_1, \cdots, \gamma_n (\in \mathfrak{F}^n$ $\exists B(\gamma_i) \geq 0$, such that for any subset $X \subseteq [n]$, we have:*

$$\sum_{\{i,j\} \subset X} V(\gamma_i, \gamma_j) \geq -\sum_{i \in X} B(\gamma_i) \tag{VII.2.22}$$

The key theorem is:

**Theorem VII.2.4** (Procacci [Pro07])**.** *For the cluster coefficients defined above, with a stable interaction, we have the following inequality:*

$$\left|\phi^T(\gamma_1, \cdots, \gamma_n)\right| \leq e^{\sum_{i=1}^{n} B(\gamma_i)} \sum_{\tau \in \mathfrak{a}[n]} \prod_{\{i,j\} \in E(\tau)} F(\gamma_i, \gamma_j) \tag{VII.2.23}$$

*where*

$$F(\gamma_i, \gamma_j) := \begin{cases} \left|e^{-V(\gamma_i, \gamma_j)} - 1\right| & \textit{if } \gamma_i \iota \gamma_j \\ |V(\gamma_i, \gamma_j)| & \textit{otherwise} \end{cases} \tag{VII.2.24}$$

In order to prove this, it is necessary to go through modified potentials to take care of the hardcore interactions. We note also that this may be generalised to continuous systems, by realising that we may define the incompatibility condition as corresponding to two particles being closer than the hardcore radius to each other. The algebra inside of the sum goes through for the potentials depending on position points and then must be integrated after. This interpretation is given after the derivation.

For a given $H \in \mathbb{R}_+$, we have the modified potential (for polymers):

$$V_H(\gamma_i, \gamma_j) = \begin{cases} H & \text{if } \gamma_i \iota \gamma_j \\ V(\gamma_i, \gamma_j) & \text{otherwise} \end{cases} \tag{VII.2.25}$$

For large enough $H$, the stability property of $V_H$ is inherited from that of $V$.

**Lemma VII.2.5** (Proccaci)**.** *For any fixed $n \in \mathbb{N}$ and $(\gamma_1, \cdots, \gamma_n) \in \mathfrak{F}^n$ $\exists H_0$ such*

*that* $\forall H \geq H_0$ *and* $\forall X \subset [n]$*:*

$$\sum_{\{i,j\}\subset X} V_H(\gamma_i, \gamma_j) \geq -\sum_{i\in X} B(\gamma_i) \tag{VII.2.26}$$

We may then apply the tree formula to the potential $V_H$.

$$\sum_{g\in\mathcal{C}[n]} \prod_{\{i,j\}\in E(g)} \left(e^{-V_H(\gamma_i,\gamma_j)} - 1\right) = \sum_{\tau\in\mathfrak{a}[n]} \prod_{\{i,j\}\in E(\tau)} (-V_H(\gamma_i, \gamma_j) \int \mathrm{d}\mu_\tau((t)_{[n-1]}, (X)_{[n-1]})\, e^{-K_H((X)_{[n-1]},(t)_{[n-1]})} \tag{VII.2.27}$$

where $K_H((X)_{[n-1]}, (t)_{[n-1]}) := \sum_{1\leq i<j\leq n} t_1(\{i,j\}) \cdots t_{n-1}(\{i,j\}) V_H(\gamma_i, \gamma_j)$.

For fixed $\tau = ([n], E(\tau)) \in \mathfrak{a}[n]$ and $(\gamma_1, \cdots, \gamma_n) \in \mathfrak{F}^n$, we define:

$$w_N^\tau(\gamma_1, \cdots, \gamma_n) := \prod_{\{i,j\}\in E(\tau)} (V_H(\gamma_i, \gamma_j) \int \mathrm{d}\mu_\tau((t)_{[n-1]}, (X)_{[n-1]})\, e^{-K_H((X)_{[n-1]},(T)_{[n-1]})}) \tag{VII.2.28}$$

Define the set:

$$E^H(g) := \{\{i,j\} \in E(g) |\, \gamma_i \iota \gamma_j\} \tag{VII.2.29}$$

This provides a partition of the edge set $E(g)$ into the disjoint union of $E^H(g)$ and $E(g) \setminus E^H(g)$.

We also define a new modified potential:

$$U_{(1-\varepsilon)H}(\gamma_i, \gamma_j) = \begin{cases} (1-\varepsilon)H & \text{if } \gamma_i \iota \gamma_j \\ 0 & \text{otherwise} \end{cases} \tag{VII.2.30}$$

We may rewrite the factor in the exponent of (VII.2.28) as:

$$K_H((X)_{[n-1]}, (t)_{[n-1]}) = K_{U_{(1-\varepsilon)H}}((X)_{[n-1]}, (t)_{[n-1]}) + K_{V_{\varepsilon H}}((X)_{[n-1]}, (t)_{[n-1]}) \tag{VII.2.31}$$

We then observe lower bounds for these exponents, which will provide upper bounds in a tree-graph inequality. Firstly, the stability of $V$ provides:

$$K_{V_{\varepsilon H}}((X)_{[n-1]}, (t)_{[n-1]}) \geq \sum_{i=1}^n B(\gamma_i) \tag{VII.2.32}$$

We also have the bound for the other potential:

$$K_{U_{(1-\varepsilon)H}}((X)_{[n-1]},(t)_{[n-1]}) \geq \sum_{1\leq i<j\leq n} t_1(\{i,j\})\cdots t_{n-1}(\{i,j\})v^\tau_{i,j} - \left|E(\tau)\setminus E^H(\tau)\right|\eta \tag{VII.2.33}$$

where

$$v^\tau_{i,j} = \begin{cases} (1-\varepsilon)H & \text{for}\{i,j\}\in E^H(\tau) \\ \eta & \text{for } \{i,j\}\in E(\tau)\setminus E^H(\tau) \\ 0 & \text{otherwise} \end{cases} \tag{VII.2.34}$$

We thus obtain the upper bound on $w^\tau_H$:

$$\begin{aligned} w^\tau_H(\gamma_1,\cdots,\gamma_n) \leq e^{\sum_{i=1}^n B(\gamma_i)+\eta\left|E(\tau)\setminus E^H(\tau)\right|}\left(\prod_{\{i,j\}\in E(\tau)\setminus E^H(\tau)}|V(\gamma_i,\gamma_j)|\right)\left(\frac{1}{\eta}\right)^{\left|E(\tau)\setminus E^H(\tau)\right|} \\ \times\left(\frac{1}{1-\varepsilon}\right)^{\left|E^H(\tau)\right|}\prod_{\{i,j\}\in E(\tau)} v^\tau_{i,j}\left(\int \mathrm{d}\mu_\tau((t)_{[n-1]},(X)_{[n-1]})\, e^{-\sum_{1\leq i<j\leq n} t_1\cdots t_{n-1}v^\tau_{i,j}}\right) \end{aligned} \tag{VII.2.35}$$

The next stage is to apply the tree-graph identity again to $v^\tau_{i,j}$

$$\begin{aligned} &\prod_{\{i,j\}\in E(\tau)} v^\tau_{i,j}\left(\int \mathrm{d}\mu_\tau((t)_{[n-1]},(X)_{[n-1]})\, e^{-\sum_{1\leq i<j\leq n} t_1\cdots t_{n-1}v^\tau_{i,j}}\right) \\ &\qquad = \left|e^{-\eta}-1\right|^{\left|E(\tau)\setminus E^H(\tau)\right|}\prod_{\{i,j\}\in E^H(\tau)}\left|e^{-U_{(1-\varepsilon)H}(\gamma_i,\gamma_j)}-1\right| \end{aligned} \tag{VII.2.36}$$

Noting that $e^{\eta\left|E(\tau)\setminus E^H(\tau)\right|}\left|e^{-\eta}-1\right|^{\left|E(\tau)\setminus E^H(\tau)\right|} = \left|e^{\eta}-1\right|^{\left|E(\tau)\setminus E^H(\tau)\right|}$ and taking $\eta\downarrow 0$ and $\varepsilon\downarrow 0$, we achieve:

$$w^\tau_H(\gamma_1,\cdots,\gamma_n) \leq e^{\sum_{i=1}^n B(\gamma_i)}\prod_{\{i,j\}\in E^H(\tau)}\left|e^{-V(\gamma_i,\gamma_j)}-1\right|\prod_{\{i,j\}\in E(\tau)\setminus E^H(\tau)}|V(\gamma_i,\gamma_j)| \tag{VII.2.37}$$

which is the desired result as we take the limit as $H\uparrow\infty$.

For the continuous case, given the hardcore radius as $r_{\mathrm{hc}}$, we obtain the following factor for each pair $\{i,j\}$:

$$\left(\mathbb{1}_{|x_i-x_j|<r_{\mathrm{h.c.}}} + |\Phi_{i,j}|\,\mathbb{1}_{|x_i-x_j|\geq r_{\mathrm{h.c.}}}\right) \tag{VII.2.38}$$

This corresponds to recognising that the above goes through as it is if we consider the polymers as being points in space and do the above for fixed points, before

taking the integrals over space.

## VII.3 The Penrose Construction: Partitionality of Connected Graphs

The following presentation of the notion of a partition scheme can be found in [FePr07]. Given a graph $G = (V, E)$, let $\mathcal{C}_G$ denote the set of all connected spanning subgraphs of $G$ and $\mathfrak{a}_G$ the family of trees in $\mathcal{C}_G$. Define a partial order of $\mathcal{C}_G$ by bond inclusion: $G \leq \tilde{G} \iff E(G) \subset E(\tilde{G})$. For $G \leq H$, we define the set $[G, H] = \{K | G \leq K \leq H\}$ The Penrose construction partitions the set of connected graphs into subsets of the form $[T, R(T)]$, where $R : \mathfrak{a}_G \to \mathcal{C}_G$. Many different constructions can be used to achieve an $R$. Penrose gave one explicit example in [Pen67].

**Definition** (Partition Scheme)**.** *A partition scheme for a family $\mathcal{C}_G$ of connected graphs is any map $R : \mathfrak{a}_G \to \mathcal{C}_G \ \tau \mapsto R(\tau)$, such that:*

*i) $E(R(\tau)) \supset E(\tau)$ and*

*ii) $\mathcal{C}_G$ is the disjoint union of the sets $[\tau, R(\tau)]$ for $\tau \in \mathfrak{a}_G$.*

The Penrose scheme is as follows:

Fix an enumeration $v_0 \cdots v_n$ for the vertices of $G$ and for each $\tau \in \mathfrak{a}_G$ (which we view as a tree rooted at $v_0$). For any vertex $v_i$ of $\tau$, let $d(i)$ be the tree distance of the vertex $v_i$ to $v_0$ and let $v_{i'}$ be the predecessor of $v_i$ i.e. $d(i') = d(i) - 1$ and $\{v_{i'}, v_i\} \in E(\tau)$. We associate to $\tau$, the graph $R_{\text{Pen}}(\tau)$ found by adding (only once) to $\tau$ all edges $\{v_i, v_j\} \in E(\tau)$ such that either:

P1 $d(i) = d(j)$ edges between vertices at same generation

P2 $d(j) = d(i) - 1$ and $i' < j$ edges between vertices one generation away

For a partition scheme $R$, denote by $\mathfrak{a}_R := \{\tau \in a_G | R(\tau) = \tau\}$ the set of $R$-trees. In particular, $\mathfrak{a}_{R_{\text{Pen}}}$ is the set of Penrose trees. We have the following proposition:

**Proposition VII.3.1** (Alternating Connected Graph Sum)**.**

$$\sum_{g \in \mathcal{C}_G} (-1)^{|E(g)|} = (-1)^{|V|-1} |\mathfrak{a}_R| \tag{VII.3.1}$$

*for any partition scheme $R$.*

*Proof.* For any numbers $(x_e)_{e\in E}$, we have :

$$\begin{aligned}\sum_{g\in\mathcal{C}_G}\prod_{e\in E(G)}x_e &= \sum_{\tau\in\mathfrak{a}_G}\prod_{e\in E(\tau)}x_e\sum_{F\subset E(R(\tau))\setminus E(\tau)}\prod_{e\in F}x_e\\ &= \sum_{\tau\in\mathfrak{a}_G}\prod_{e\in E(\tau)}x_e\prod_{e\in E(R(\tau))\setminus E(\tau)}(1+x_e)\end{aligned}\tag{VII.3.2}$$

and for any tree $|E(\tau)| = |V| - 1$. □

In models where we only have soft repulsion, only $|1+x_e| \leq 1$ is guaranteed. This then gives the bound:

$$\left|\sum_{G\in\mathcal{C}_G}\prod_{e\in E(G)}x_e\right| \leq \sum_{\tau\in\mathfrak{a}_G}\prod_{e\in E(\tau)}|x_e| \leq |\mathfrak{a}_G|\tag{VII.3.3}$$

It is necessary to remark here that the identity (VII.3.1) is analogous to that obtained in Chapter III through the paper of Bernardi [Ber08] and noticed by Ducharme Labelle and Leroux [DLL07] for $G = K_n$, the complete graph on $n$ vertices. Indeed, the Penrose construction should allow one to create an involution like [Ber08] in order to find the appropriate combinatorial factor $(n-1)!$ from the trees which are fixed under the Penrose map $R_{\text{Pen}}$.

We see that considering the tree as being rooted at $v_0$, we are required to have precisely one vertex in a generation. This necessarily gives a linear tree. If we identify $v_0$ with the vertex labelled 1, then we know that the position of vertex 1 is decided already and we have to determine the positions of $i \in [2, n]$, which are defined uniquely by their distance from 1, which corresponds to a bijection, $f : [2, n] \to [n-1]$, giving the $(n-1)!$ factor.

**The connection between Partitions and Involutions as in Chapter III**

We let $d_1^g(i)$ denote the graph distance from the vertex labelled 1 to the vertex labelled $i$.

**Definition** (Penrose Active Edges)**.** *An edge $\{i, j\}$ is called* Penrose active *for $g$ if, either:*

*i)* $d_1^g(i) = d_1^g(j)$ *or*

*ii)* $d_1^g(i) = d_1^g(j) + 1$ *and* $\exists i' < j$ *such that* $\{i, i'\} \in E(g)$ *with* $d_1^g(i') = g_1^g(j)$.

We let $e^\star_{\text{Pen},g}$ be the greatest Penrose active edge for $g$ in lexicographic order. The mapping $\Psi : g \mapsto g \oplus e^\star_{\text{Pen},g}$ is an involution on connected graphs.

We can also go the other way and find a Bernardi construction to deliver an appropriate partition. The map $R : T_G \to C_G$, which adds to $\tau$ all externally active edges for the given tree graph $\tau$ is the appropriate partition scheme.

Recall that an edge $\{i, j\}$ is externally active by Bernardi if there is a path from $i$ to $j$ in the graph, using edges all larger than $\{i, j\}$ in the lexicographic order. When we have a connected graph, we remove the largest present externally active edge each time, since each connected graph that is not a tree will have a present externally active edge. This gives a unique tree for each connected graph as there is only ever one edge to remove at each stage. We also note that each edge that has been removed retains its status as being externally active, due to the fact that no larger edge is removed at any subsequent stage and so the collection of the externally active edges at the final stage corresponds directly with the possible connected graphs to make from it.The graph is therefore able to be attained the other way.

A more general partitioning construction is possible for matroids. This is presented in section VII.4. We see that Bernardi gives a special case of this, but that the Penrose construction cannot be realised as part of this.

## VII.4 Extension to Matroids

A generalisation of the conditions we have for our connected graphs to generate the tree identities comes from the concept of matroids. The connection is presented in the work of Faris [Far12a] and offers a framework in which one can view the main concepts and understand the scope of the approach. This section begins with the basic definition of matroids and then indicates how the tree-graph identitiy can be applied in this case through the analogue of the fundamental theorem of calculus of a matroid. The motivation is that this should be further generalised to partially ordered sets in order to provide an idea for two-connected graphs.

### VII.4.1 Definitions

A matroid is the generalisation of the notion of linear independence of a vector space. The definitions below capture the main algebraic aspects of this notion of linear independence and they can be generalised as a structure in many other contexts. The key application to be considered here is the application to graphs, the so-called graphic matroids.

**Definition.** *A matroid $\mathcal{M}$ on $E$ is given by a subset $\mathcal{I} \subseteq \mathcal{P}(E)$. An element of $\mathcal{I}$ is called an independent set.*

*The independent sets obey the following axioms:*

1. $\emptyset \in \mathcal{I}$ *(non empty)*
2. *If* $X \in \mathcal{I}$ *and* $X' \subset X$ *then* $X' \in \mathcal{I}$ *(downward closed)*
3. *If* $X \in \mathcal{I}$ *and* $Y \in \mathcal{I}$ *and* $|X| < |Y|$*, then there exists* $l \in Y \setminus X$ *with* $X \cup \{l\} \in \mathcal{I}$ *(augmentation property)*

Given a matroid $\mathcal{M}$, we denote the ground set by $E(\mathcal{M})$ and the collection of independent sets by $\mathcal{I}(\mathcal{M})$. For a graphic matroid, where the ground set is the collection of edges, forests are the independent sets. These obey the three axioms of independent sets. Certainly the graph with no edges is considered a forest (we have no cycles). If we consider the subgraph of any forest, we get a forest (downward closed). The trickier property to understand is the augmentation property. If we consider two forests, one with less edges than the other, we must choose an edge from the larger forest to place in the smaller forest that retains the acyclic property. This is always possible by realising that the number of connected components in the larger forest is smaller than the number of connected components in the smaller forest. We therefore have an edge in the larger forest that would connect two connected components in the smaller forest. This edge will therefore not create a cycle.

**Definition.** *A maximal independent set* $X \in \mathcal{I}(\mathcal{M})$ *is called a basis. The set of bases is denoted* $\mathcal{B}(\mathcal{M})$*.*

For a graph $g$, the matroid structure consists of all possible subgraphs. The basis of this matroid is comprised of forests, which have the same connected component decomposition as $g$.

**Definition.** *The rank of a matroid* $\mathcal{M}$*,* $\mathrm{rk}(\mathcal{M})$ *is the cardinality of a basis element.*

All bases have the same cardinality and so the rank is well defined. A matroid can be defined by its set of bases, since $X \in \mathcal{I}(\mathcal{M})$ if and only if $X \subseteq Y$, for some $Y \in \mathcal{B}(\mathcal{M})$.

**Definition.** $Y$ *is a spanning set for the matroid, if there exists a basis element* $X \in \mathcal{B}(\mathcal{M})$ *such that* $Y \supseteq X$*.*

**Definition** (Restricted Matroid and Rank)**.** *Given a matroid,* $\mathcal{M}$*, consider* $X \subseteq E(\mathcal{M})$*. There is a matroid* $\mathcal{M}|_X$*, which is the restriction of* $\mathcal{M}$ *to* $X$*. It has ground set* $X$ *and* $\mathcal{I}(\mathcal{M}|_X) = \{Y \in \mathcal{I}(\mathcal{M}) | Y \subseteq X\}$*.*

*For $X \subseteq E(\mathcal{M})$, the rank of $X$, $\mathrm{rk}(X)$ is the rank of the matroid $\mathcal{M}|_X$ or alternatively the cardinality of the largest independent subset of $X$.*

We note that $\mathrm{rk}(X) = |X|$ if and only if $X$ is independent, so the rank function completely determines the matroid.

**Definition** (Loops and Loopless Matroids)**.** *A set $X$ is dependent if it is not independent. A loop is an element $l$ such that $\{l\}$ is dependent. A loopless matroid is one where all one element subsets are independent.*

**Definition.** *Suppose $X$ is an independent set, so its rank $\mathrm{rk}(X) = |X|$. Span$(X)$ consists of all $l \in E$ such that $X \cup \{l\}$ also has rank $\mathrm{rk}(X)$.*

**Lemma VII.4.1.** *Let $X$ be an independent set. Suppose that $l$ is in the span of $X$. Then the collection of sets $Z \subseteq X$ such that $l$ is in the span of $Z$ has a least element. This element is denoted by $S(l, X)$.*

**Remark 39.** *The set $S(l, X)$ is characterised by the property that, for $Z \subseteq X$, we have $S(l, X) \subseteq Z$ if and only if $l$ is in the span of $Z$.*

### VII.4.2 Statement of the Fundamental Theorem of Calculus on Matroids

The fundamental theorem of calculus can be extended from that for graphs to the more general matroids. An important point to consider is what sort of structure on a partially ordered set will allow for a generalisation of this fundamental theorem of calculus. Given variables $(s_l)_{l\in E}$ and for a given $X \in \mathcal{I}$, the corresponding variables $(u_l)_{l\in X}$, we define $s_X(u)_l$ for $l$ in $E$ analogously to (VII.1.3):

1. If $l \in \mathrm{Span}(X)$, then $s_X(u)_l = \min_{e\in S(l,X)} u_e$
2. If $l \notin \mathrm{Span}(X)$, then $s_X(u)_l = 0$

**Remark 40** (Loops)**.** *If $l$ is a loop, then for every $X$, we have $l \in$ Span$(X)$ and so $S(l, X) = \emptyset$. $s_X(u)_l = 1$ in this case. For $X = \emptyset$ Span$(\emptyset)$ is the set of loops and so $s_\emptyset(u)_l = 1$ if $l$ is a loop and $0$ otherwise.*

**Theorem VII.4.2** (Fundamental Theorem of Calculus on Matroids)**.** *Consider a matroid $M$ with ground set $E$ and with a collection $\mathcal{I} \subseteq \mathcal{P}(E)$ of independent sets, Let $f$ be a smooth function of variables $s_l$ for $l \in E$. Then*

$$f(1) = \sum_{X\in\mathcal{I}} \int_{[0,1]^X} \left(\prod_{l\in X} \frac{\partial}{\partial s_l} f\right)(s_X(u)) \prod_{l\in X} \mathrm{d}u_l \tag{VII.4.1}$$

We can recognise the relationship with the forest formula of Brydges Kennedy, Abdesselam and Rivasseau, (VII.1.28), by using the interpretation of the graphic matroid. We obtain the forest formula by setting $f((u_e)_{e\in E}) = \exp(\sum_{e\in E} u_e \Phi_e)$.

To make the connection to statistical mechanics, consider the case of *product functions*, which are of the form:

$$f_X(s) = \prod_{l\in X} f_l(s_l) \tag{VII.4.2}$$

We take the variables $s_l \in [0,1]$ and call $f_l$ normalised if $f_l(0) = 1$.

$f$ may also be expressed in distributive law form:

$$f_X(s) = \prod_{l\in X} (1 + s_l t_l) \tag{VII.4.3}$$

The $Y \subseteq X$ partial derivative is $\left(\prod_{l\in Y} \frac{\partial}{\partial s_l}\right) f_X(s) = \prod_{l\in Y} t_l f_{X\setminus Y}(s)$. If we have $-1 \leq t_l \leq 0$, then $0 \leq 1 + s_l t_l \leq 1$ and $0 \leq f_{X\setminus Y}(s) \leq 1$.

In *exponential form*, we may write our function as:

$$f_X(s) = \prod_{l\in X} \exp(-s_l v_l) \tag{VII.4.4}$$

We obtain for the $Y \subseteq X$ partial derivative:

$$\left(\prod_{l\in Y} \frac{\partial}{\partial s_l}\right) f_X(s) = \prod_{l\in Y} (-v_l) f_X(s) \tag{VII.4.5}$$

If $v_l \geq 0$, then $0 \leq \exp(-s_l v_l) \leq 1$ and hence $0 \leq f_X(s) \leq 1$.

The distributive law form reflects the Mayer $f$ function approach to cluster and virial expansions, whereas using the exponential straight away echoes the approach of Abdesselam and Rivasseau [AbRi94].

In extending bounds from positive potentials to more general potentials, it is necessary to use stability in the case of connected graphs, to bound the extra edge factors in trees. A generalisation of this also exists for matroids.

**Definition** (Flats)**.** *For $X \subseteq E(\mathcal{M})$ arbitrary subset. Span$(X)$, written cl$(X)$ the closure of $X$ is the greatest set $Y$ such that $X \subseteq Y$ and* $\mathrm{rk}(X) = \mathrm{rk}(Y)$*. A subset $F$ is* flat *if Span$(F) = F$. This is the same as saying that $F$ is maximal with respect to having rank* $\mathrm{rk}(F)$*. The collection of flats $\mathcal{F}(\mathcal{M})$ determines uniquely the matroid $\mathcal{M}$. Considering the set $\mathcal{F}(\mathcal{M}) \setminus \{E(\mathcal{M})\}$, we can say that a set is non-spanning if*

*and only if it is a subset of a set in this collection.*

For the graphic matroid, a flat is a graph all of whose connected components are complete graphs.

**Theorem VII.4.3** (Stability for Matroids)**.** *For a matroid $\mathcal{M}$ with ground set $E(\mathcal{M})$, if $\exists c \geq 0$ constant such that for each flat $Y \subset E(\mathcal{M})$, the interaction satisfies:*

$$\sum_{l \in Y} \Phi_l \geq -c \tag{VII.4.6}$$

*then for an independent set $I$, with fixed variables $(u_l)_{l \in I}$, the interpolated interaction satisfies the bound:*

$$\sum_{l \in E} s_I(u)_l \Phi_l \geq -c \tag{VII.4.7}$$

**Lemma VII.4.4.** *$I$ independent set in a matroid $\mathcal{M}$, $Y$ =Span($I$), fix variables $(u_l)_{l \in I}$, and let $0 \leq \tau \leq 1$. $Y(\tau) = \{l \in E | s_I(u)_l \geq \tau\}$ is a subset of $Y$ and is a flat.*

**Lemma VII.4.5.** *$I \in \mathcal{I}(M)$ and $Y$ =Span($I$) and fix $(u_l)_{l \in I}$, then for each $0 \leq \tau \leq 1$, there is a flat $Y(\tau) \subseteq Y$ such that:*

$$\sum_{l \in E} s_I(u)_l \Phi_l = \int_0^1 \sum_{l \in Y(\tau)} \Phi_l \mathrm{d}\tau \tag{VII.4.8}$$

**Theorem VII.4.6** (Fundamental Theorem of Calculus for Spanning Set)**.** *For every normalised product function*

$$\sum_{S \in S(M)} \prod_{l \in S} (f_l(1) - 1) = \sum_{B \in \mathcal{B}(M)} K_B(f_{E(M)}) \tag{VII.4.9}$$

where we define, for $I$ an independent set and $I \subset \mathrm{Span}(T) \subset X$:

$$K_I(f_X) := \int_{[0,1]^I} \left( \prod_{l \in I} \frac{\partial}{\partial s_l} f_X \right) (s_I(u)) \prod_{l \in I} \mathrm{d}u_l \tag{VII.4.10}$$

and note that $K_I(f_X) = K_I(f_{\mathrm{Span}(I)})$, where $f$ is a normalised product functions. That is $f_l(0) = 1 \; \forall l \in E$.

The above theorem is proved using the poset structure of the collection of flats with inclusion as the order: $(\mathcal{F}, \subset)$. One obtains the Möbius inverse of the function $\phi(X) := f_X(1)$ in two different ways. The first way emulates the method

of Mayer $f$ functions, where one writes:

$$f_X(1) = \sum_{S \subset X} \prod_{l \in S} (f_l(1) - 1) \tag{VII.4.11}$$

and through using the fundamental theorem of calculus. By uniqueness of Möbius inverse, one equates the two inverses achieved through the two methods.

Applying the theorem in the case of distributive law functions, one has:

**Corollary VII.4.7.** *Given the collection of variables $(t_l)_{l \in E}$, satisfying $-1 \leq t_l \leq 0$, we have:*

$$\left| \sum_{S \in S(M)} \prod_{l \in S} t_l \right| \leq \sum_{B \in \mathcal{B}(M)} \prod_{l \in B} |t_l| \tag{VII.4.12}$$

Realising the connection to the graphic matroid, through the fact that spanning sets are all connected graphs and the basis elements are trees. This gives the Penrose inequality.

We may also apply the theorem to the exponential functions to obtain:

**Corollary VII.4.8.** *For a matroid $\mathcal{M}$ and $\forall l \in E$, we fix numbers $\Phi_l$, such that $\exists c \geq 0$ constant such that for every flat $F$ of $M$, there is a stability bound $\sum_{l \in F} \Phi_l \geq -c$, then:*

$$\left| \sum_{S \in S(\mathcal{M})} \prod_{l \in S} (e^{-\Phi_l} - 1) \right| \leq e^c \sum_{B \in \mathcal{B}(\mathcal{M})} \prod_{l \in B} |\Phi_l| \tag{VII.4.13}$$

In this case, we obtain the Brydges and Federbush Graph Tree inequality for the graphic matroid.

### VII.4.3 Internally and Externally Active Edges

In Sokal's paper [Sok05], there is a very helpful generalisation of the Penrose partition to matroids.

For $R \subseteq S \subseteq E$, we define $[R, S] = \{A | R \subseteq A \subseteq S\}$. We call $[R, S]$ a *molecule* if we may express $S$ as a disjoint union $S = R \cup F \cup T$, such that for each $A \in [R, S]$ $\mathrm{rk}(A) = \mathrm{rk}(R) + |A \cap F|$. We note that $[R, S]$ a molecule implies that $[R', S'] \subseteq [R, S]$ is a molecule.This concept is a generalisation of the Penrose partition.

**Definition** (Dual of a Matroid)**.** *The dual of a matroid is defined on the same ground set, but has a dual rank function $rk^\star$, defined by:*

$$rk^\star(A) := |A| - rk(E) + rk(E \setminus A) \tag{VII.4.14}$$

Let $\mathcal{B}$ be the set of bases for $E$. The dual basis set is then $\mathcal{B}^\star = \{E\backslash B | B \in \mathcal{B}\}$. We fix a total order on $E$ in the following.

**Definition** (Externally Active)**.** *Let $B \in \mathcal{B}$. An element $e \in E \setminus B$ is* externally active *on $B$ if $e$ is dependent on the list of elements of $B$ larger than it. We let $E(B)$ be the set of externally active elements with cardinality $e(B)$, which is called the external activity of $B$.*

**Definition** (Internally Active)**.** *An element $e \in B$ is* internally active *on $B$, if in the dual matroid $e$ is externally active on the complement $B^c = E \setminus B \in \mathcal{B}^\star$. We denote by $I(B) = E^\star(B^c)$ the set of internally active elements with cardinality $i(B)$, which is called the internal activity of $B$.*

**Proposition VII.4.9.** *$2^E$ can be written as the disjoint union:*

$$2^E = \bigsqcup_{B\in\mathcal{B}} [B \setminus I(B), B \cup E(B)] \tag{VII.4.15}$$

*for each $B \in \mathcal{B}$ $[B \setminus I(B), B \cup E(B)]$ is a molecule with $F = I(B)$ and $T = E(B)$.*

For the case of the graphical matroid, we recall that the bases are the collection of trees. If we use the lexicographical order on the edges, then an edge is externally active for a tree $\tau$ in this sense, if and only if it is externally active in the sense of Bernardi [Ber08]. This is due to the fact that all independent sets are forests and so a set of edges is dependent if it creates a cycle. We emphasise that for connected graphs, internally active edges play no role, since trees are minimally connected graphs.

We note that the Penrose construction does not fit in the construction given above. In Figure VII.2, we see that we would add the dashed edge in each case. In order to do this, we cannot have a consistent ordering on the edges $\{2, 3\}$, $\{2.4\}$ and $\{3, 4\}$.

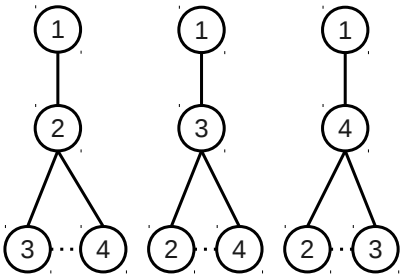



Figure VII.2: Three Graphs to Indicate that the Penrose Construction is Different

The key conclusion of these observations is that we are able to modify these constructions towards two-connected graphs, where the notion of internally active will be important.

## VII.5 Extensions to the Tree-Graph Inequalities: Kruskal's Algorithm and Edge Ordering

Related ideas to extending and improving the tree graph identities can be found in the work of Helmuth [Hel14] and Rivasseau and Tanasa [RiTa14]. The key idea is to fully understand how to use a tree-graph identity that respects particular features of the potential to ensure the expansion converges.

The first key idea is to totally order the edges. This has already been used to define the concept of externally and internally active edges. In the interpretation of these two papers, the ordering allows Kruskal's algorithm to run and choose a spanning tree. To symmetrise over the orderings, a variable $s_{i,j}$ is introduced for each edge $\{i, j\}$. If we integrate each of these variables over the interval $[0, 1]$, we see that we may divide $[0, 1]^{\binom{n}{2}}$ into simplices, where a total order on the variables is defined. The application of Kruskal's algorithm within each of these sectors gives the Brydges-Kennedy-Abdesselam-Rivasseau Tree-Graph identity. In [RiTa14], these sectors of the integral are identified with Hepp sectors.

The concept of splitting the graph contribution amongst its various spanning trees is generalised in [RiTa14] as rewriting amplitudes as:

$$S = \sum_{g \in \mathcal{G}[n]} A_g = \sum_{g \in \mathcal{G}[n]} \sum_{\tau \in \mathfrak{a}[n]} \sum_{\tau \leq g} w(g, \tau) A_g = \sum_{\tau \in \mathfrak{a}[n]} A_\tau \tag{VII.5.1}$$

where $A_\tau = \sum_{g \geq \tau} w(g, \tau) A_g$. This $w(g, \tau)$ gives how the contribution of $g$ is split between the various different spanning subtrees $\tau$. The formal definitions are as follows:

**Definition** (Probability Measure on Trees)**.** *A probability measure on trees is a set of positive weights $w(g, \tau)$ for any labelled connected graph $g$ and any tree $\tau \leq g$, such that:*

$$\sum_{\tau \leq g} w(g, \tau) = 1 \tag{VII.5.2}$$

*It is called* rational *if $w(g, \tau) \in \mathbb{Q}$ $\forall g$ and $\tau$.*
*It is called* symmetric *if $w(g, \tau) = w(\sigma(g), \sigma(\tau))$ for all permutations $\sigma$ on the label set.*

In order to achieve convergence of the rearranged series, the definition of *constructive* weights is introduced:

**Definition** (Constructive Weights)**.** *A probability measure on trees is called* constructive*, if we have a $\tau$-dependent probability measure $(\Omega_\tau, \Sigma_\tau, \mu_\tau)$ and a $(\tau, u)$-dependent real positive symmetric matrix $X^\tau_{i(l),j(l)}(u)$ for every $u \in \Omega_\tau$, such that the weights can be written as:*

$$w(g,\tau) = \int_{\Omega_\tau} \mathrm{d}\mu_\tau(u) \prod_{l \in E(\tau)} X^\tau_{i(l),j(l)}(u) \tag{VII.5.3}$$

*where $i(l)$ and $j(l)$ denote the two endpoints of the edge $l$. The diagonal entries in the matrix are all* 1*.*

A special case of constructive tree weights is that of the Brydges-Kennedy-Abdesselam-Rivasseau Graph-Tree identity, where the matrix picks the smallest edge on the route in the tree between the endpoints.

## VII.6 Conclusions & Open Questions

Tree graph identities and inequalities are fundamental tools in understanding the cluster expansion. They provide a useful way to resum such series and are also found in other contexts [RiWa10a, RiWa10b]. Furthermore, they are an effective way of dealing with cancellations in the cluster expansion. A key idea, which is featured in [BrMa14], is the fact that the tree identity provides a useful majorant for the cluster expansion. These are also linked to the combinatorics found for the two extreme cases for positive potentials of the Tonks gas and the one-particle hardcore gas as explained by Bernardi [Ber08].

The concept of matroids, brought to the attention of cluster expansions, through Faris [Far12a], but also used in the work of Sokal [Sok05] for Potts models, provides the key generalisation of the tree graph identities to more general structures. The key idea is in an effective reorganisation of series expansions, so that cancellations are dealt with effectively. It is the first step in providing further generalisation, such as that required for two-connected graphs.

Unfortunately, two-connected graphs do not form a natural matroid structure in order to apply these more generalised techniques to and the attempt to understand the two-connected graphs in terms of trees, requires a more in depth understanding of the modified tree weights. The combinatorial factors need to be udnerstood better, or interpreted in a more practical way. Furthermore, the combinatorial

explanation provided in Chapter III needs to be developed into a useful way of understanding the cancellations, as has been done for connected graphs.

The key to understanding and developing these ideas further, will be to fully see how far this fundamental theorem of calculus can be generalised and what features a poset will require. A generalisation to a wider class of posets would be interesting and useful for the virial expansion if it may include the poset of two-connected graphs.

# Chapter VIII

# Estimation Methods, Minimal two-connected Graphs and Graph Partitions

In [Gro67], Groeneveld gives some ideas about how to move forward from our knowledge of the upper and lower bounds in Section I.7. The main idea comes from the success of the past attempts to obtain good estimates of the expansions afforded to us by Mayer, from the derivations made in Chapters I and II. The configuration integrals are represented by graphs of different types in each case. However, the configuration integrals are hard to compute and the evaluation of these integrals has only been done in a limited number of special cases. The problem is then moved to creating a reasonable approximation method for the given integrals based on those integrals we can compute. The approach of Lieb [Lieb63], Lebowitz and Percus [LePer63] and Penrose [Pen63b], rely on estimating remainders from taking a truncated series, whereas Groeneveld indicates that the better approach is to create an estimation method to give a consistent set of upper and lower bounds. This chapter aims to follow these ideas of Groeneveld, presented in Section VIII.1, through various estimation schemes and obtain such a scheme for two-connected graphs.

The approach of this chapter is to understand an analogue of trees for two-connected graphs. We find that flower graphs, which are defined in Section VIII.2, are the appropriate idea for minimal two-connected graphs. Howevever, the partitioning doesn't work in the same way. Furthermore, the number of such graphs is too large to allow us to bound a resummation in terms of such graphs.

Some approximation schemes have already been given such as through the Percus Yevick equations [PeYe58, Ste63] and their relationship with series-parallel

graphs. There is also the Born-Green-Yvon treatment [BoGr49] and the approach through Mayer Montroll equations [MMon41], which are used as appropriate approximations. The papers of van Leeuwen, Groeneveld and de Boer [LGB59,BLG64] explain the connections of these virial equations of state and the errors made in the approximations.

Furthermore, the Ree Hoover expansion, first introduced in [RH64b], reduces the number of graphs needed to be integrated and is applied to the case of the hard-core gas with success in estimating virial coefficients up to the tenth order through Monte Carlo simulations in the work of McCoy and Clisby [ClMc06, ClMc05]. The involution defined in Chapter III also indicates some aspect on the combinatorics of this expansion. This is shown in Section VIII.3.

## VIII.1 Groeneveld's Formulation of Graphical Approximations

The key idea is that, in the various thermodynamic expansions, we have graphs of a specific type describing the configuration integral. The types of graphs can be classified by their minimal connection number $d$, and we denote the set of all such graphs by $\Gamma_d$. An estimation method, involves taking a simplified version of the graph and understanding the coefficients given by these graphs and what sort of error we have made in the simplification.

Groeneveld, in [Gro67], conveyed this method of approximation through a table, which indicates the various models in statistical mechanics and the relevant type of graph with the appropriate approximations made.

Table VIII.1: Classification of Models in Statistical Mechanics

| d | $\Gamma_d$ | $\Sigma_d$ | Expansion Parameter | Theory |
|---|---|---|---|---|
| 0 | Simple Graphs | Empty Graphs | $z$ | |
| 1 | Connected Graphs | Cayley Trees | $z$ | 1st Mayer Theorem |
| 2 | Two-connected Graphs | ? | $\rho$ | 2nd Mayer Theorem |
| 3 | Triply Connected Graphs | ? | | Series Parallel Graphs (HNC Theory) |

For $d = 0$, $\Lambda = \Sigma_0$ is the ideal gas law; for $d = 1$, $\Lambda = \Sigma_1$ gives the Cayley tree approximation; and the two methods for $d = 2$ are $\Lambda = \Lambda_{\mathrm{PY}}$ the Percus Yevick approximation and $\Lambda = \Lambda_{\mathrm{SP}}$, the series parallel or hypernetted chain (HNC) approximation.

The general method of approximation of the corresponding theory comprises of:

- A method of simplification of graphs $G \in \Gamma$ i.e. a mapping for each graph G to an integrable simplification $H$
- Some form of estimation of the edges we do not integrate over through the Mean Value Theorem of integral calculus

The integrands can be formulated as the product of $f$-functions $f_{i,j}$ and $h$-functions $h_{i,j} = 1 + f_{i,j}$, which is the main idea used in the expansions of Ree Hoover for the virial expansion.

Formally we define an $f - h$ expansion *scheme* $(d, \Lambda, \xi, \eta)$ as consisting of:

- The degree $d$ of the expansions
- $\xi$ - a partition of the Mayer graphs of the corresponding Mayer expansion, giving the $f - h$ expansion
- A collection of skeleton graphs $\Lambda \subset \Gamma_{d^\star}$ $d^\star \leq d$- which involve only $f$-lines - these are the graphs we compute numerically and $(d^\star, \Lambda)$ is called the approximation method
- $\eta$ is the simplifying mapping from $\Gamma_d$ to $\Lambda$

## VIII.2 Flower Graphs

There is a characterisation of minimal two-connected graphs given by Plummer [Plu68], which is:

**Theorem VIII.2.1** (Plummer's Theorem)**.** *Let $G$ be a two-connected graph, then $G$ is minimally two-connected, if and only if, either:*

*i) $G$ is a cycle*

*ii) If $S$ denotes the set of vertices of degree two in $G$, then there are at least two components of $G \setminus S$ and each component of $G \setminus S$ is a tree and if $C$ is any cycle in $G$ and $T$ is any component of $G \setminus S$, then the graph $(V(C) \cap V(T), E(C) \cap E(T))$ is empty or connected.*

This section proposes an alternative characterisation of minimal two-connected graphs, which are tractable to enumeration and understanding whether it is possible to obtain an analogue to the Penrose tree construction.

### VIII.2.1 Definitions

**Definition.** *A labelled two-connected graph $g$ is called a* flower graph *if we can construct it in the following (inductive) manner:*

*We start with a vertex set $V = [N]$*

*We start with an initial cycle containing* 1*. Let $S$ be the set containing the points in the cycle. We can continue inductively by repeatedly applying the next two steps until $S = V$*

*i) Consider the element $s = \min(V \setminus S)$ and construct a chain containing this element and with endpoints in $S$ and all other elements outside of $S$.*

*ii) Now redefine $S$ as $S \cup C$, where $C$ is the set of all elements in the newly formed chain.*

**Remark 41.** *We note that step i) is always possible, since our graph is two-connected, we have two distinct paths from one point to another and so two distinct paths from* 1 *to the new element to be added. These two paths will include some edges outside of $S$, because they are distinct (have no edges in common), this gives us our chain.*

First of all we would like to understand a way of partitioning the two-connected graphs according to these minimal 'flower' graphs. We formulate the validity of the partition as the following theorem:

**Theorem VIII.2.2** (The validity of a partition corresponding to flower graphs)**.** *Every two-connected graph has a unique flower graph associated to it.*

In order to prove this we first understand that we always have a cycle containing 1 and then define which cycle we will take and make sure the subsequent steps on how we build up our flower graph from a given two-connected graph is unique.

*The validity of a partition corresponding to flower graphs.* If 1 is connected to only one other vertex then removing this vertex would disconnect 1 from the rest of the graph and so either 1 has at least two neighbours or the graph is just $K_2$, which is the flower graph on two points. If the vertex labelled 1 and its incident edges are removed the graphs is still connected. There is then a path between the two neighbours of 1 and adding in the edges to 1 gives a cycle containing 1. We then choose the shortest cycle and, from the set of shortest cycles, the one with the smallest sum of the labels of vertices in the cycle. This gives a unique first cycle

We now have a set $S$, which is just the members of the cycle. Then we consider $s = \min(V \setminus S)$, where $V$ is the vertex set of the graph. Since the graph is two-connected we have two disjoint paths from $s$ to the set $S$ and so joining these two paths together we have a chain starting and ending at $S$ and going through $s$. We add the shortest chain and that with least total sum and making sure that if there is a choice, the lowest element in $S$ connects to the lowest endpoint in the chain. We now redefine $S = S \cup C$, where $C$ is the set of points in the chain. Since our set increases in cardinality by at least one each time, eventually $S = V$, since $V$ is finite. □

### VIII.2.2 The Partition

In the example of the Penrose construction in Section VII.3, it is possible to define a partition of connected graphs with respect to trees. It also has the additional property that we can obtain a maximal connected graph from a given labelled tree. The partition is then that we associate to a tree all graphs with any number of the additional edges present in the maximal graph. The advantage of this is that each additional edge $e$ has the weight $1 + f_e$, which can be bounded by 1 for positive potentials and thus not contribute.

The issue with this flower graph construction is that we cannot achieve a maximal two-connected graph in all cases. When we have a collection of possible edges to add to our flower graph, we haven't got the possibility of taking the extra edge or not as some sets of edges when taken together will give a different flower graph when we run the procedure described in Section VIII.2.1. In Figure VIII.1, we have a minimal two-connected graph denoted by the undashed lines. Adding any of the two dashed lines we have a two-connected graph which will reduce to the minimal graph through the procedure in Section VIII.2.1. However, the graph displayed in Figure VIII.2 is what the graph with both dashed lines will reduce to using the method in Section VIII.2.1.

The above construction still defines a partition since achieving the minimal flower graph is unique, but there is no straightforward way of characterising what the partition sets are like which would allow us to obtain an estimate on the error of approximating these graphs by flower graphs. This makes it inherently more difficult to define the 'estimation factor'.

### VIII.2.3 Enumeration

Let $F_n$ denote the number of flower graphs on $n$ vertices.

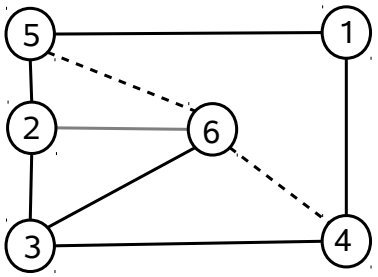



Figure VIII.1: Flower graph counterexample with two (dashed) edges that may be added individually but not together

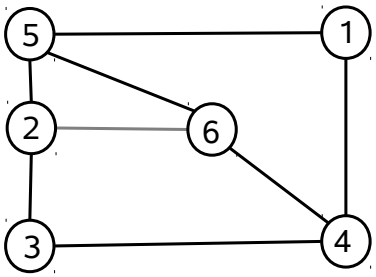



Figure VIII.2: The minimal 2-connected graph we get if both edges are added in Figure VIII.1

If we have a flower graph on $n$ vertices and we let $l_1$ be the length of the first cycle and $l_i$ be the number of elements added on the $i$th successive chain, then we can enumerate flower graphs by the number of chains/cycles we have. The first cycle has to have at least three elements for it to be a cycle. This means we have at most $n-2$ chains. We also have the restriction that $\sum_{i=1}^{k} l_i = n$. For the first cycle we have to pick the $l_1 - 1$ vertices (other than 1) for the first cycle from the $n-1$ vertices. Once the labels are chosen we can arrange them in any order in the cycle, but we have symmetry from reversing the order of the cycle about 1 and so we get a factor of $\frac{1}{2}(l_1-1)!$. We have the smallest vertex not in the cycle for the second chain, so we must choose the remaining $l_2 - 1$ from $n - l_1 - 1$. We can permute the elements in the chain as we like and we need to choose two distinct points in the cycle for endpoints of the chain, giving the factor $\binom{l_1}{2} l_2!$. The further terms follow inductively.

$$F_n \leq \sum_{k=1}^{n-1} \sum_{l_1 \geq 3} \sum_{l_2,\cdots,l_k \geq 1} \binom{n-1}{l_1-1} \binom{n-l_1-1}{l_2-1} \cdots \binom{n-l_1-\cdots-l_{k-1}-1}{l_k-1} \frac{1}{2}(l_1-1)! \binom{l_1}{2} l_2! \binom{l_1+l_2}{2} l_3! \cdots \binom{l_1+\cdots+l_{k-1}}{2} l_k! \quad \text{(VIII.2.1)}$$

$$F_n \leq \sum_{k=1}^{n-2} \sum_{l_1 \geq 3} \sum_{l_2,\cdots,l_k \geq 1} \frac{(n-1)!}{(l_1-1)!(n-l_1)!} \frac{(n-l_1-1)!}{(l_2-1)!(n-l_1-l_2)!} \cdots \frac{n-l_1-\cdots-(l_{k-1}-1)!}{(l_k-1)!(n-l_1-\cdots-l_k)!}$$
$$\frac{1}{2}(l_1-1)!l_2!\cdots l_k! \frac{l_1(l_1-1)}{2} \frac{(l_1+l_2)(l_1+l_2-1)}{2} \cdots$$
$$\frac{(l_1+\cdots+l_{k-1})(l_1+\cdots+l_{k-1}-1)!}{2} \tag{VIII.2.2}$$

We use $(l_1+\cdots l_2)^2$ instead of $(l_1+\cdots+l_s)(l_1+\cdots+l_s-1)$ since it is an upper bound and gives a simpler expression.

$$F_n \leq \sum_{k=1}^{n-1} 2^{-k} \sum_{l_1 \geq 3} \sum_{l_2,\cdots,l_k \geq 1} (n-1)! \frac{1}{n-l_1} \cdots \frac{1}{n-l_1-\cdots-l_k} l_2 \cdots l_k$$
$$l_1^2 (l_1+l_2)^2 \cdots (l_1+\cdots+l_{k-1})^2 \tag{VIII.2.3}$$

$$F_n \leq (n-1)! \sum_{k=1}^{n-2} 2^{-k} \sum_{l_1+\cdots+l_k=n} \frac{l_2 \cdots l_k l_1^2 (l_1+l_2)^2 \cdots (l_1+\cdots+l_{k-1})^2}{(n-l_1)\cdots(n-l_1-\cdots l_{k-1})} \tag{VIII.2.4}$$

We make the change of variables $t_s = l_1+\cdots+l_s$ and $t_k = n$ and we have strict inequalities between the $t_i$s, since each $l_i \geq 1$. We make the standard integral approximation for large $n$:

$$\sum_{1<t_1<t_2<\cdots<t_{k-1}<n} \frac{t_1^2 t_2^2 \cdots t_{k-1}^2}{(n-t_1)\cdots(n-t_{k-1})} (t_2-t_1)\cdots(n-t_{k-1})$$
$$= n^{2(k-1)} \sum_{0<\frac{t_1}{n}<\cdots<\frac{t_{k-1}}{n}<1} \frac{\left(\frac{t_1}{n}\right)^2 \cdots \left(\frac{t_{k-1}}{n}\right)^2}{\left(1-\frac{t_1}{n}\right)\cdots\left(1-\frac{t_{k-1}}{n}\right)} \left(\frac{t_2-t_1}{n}\right)\cdots\left(1-\frac{t_{k-1}}{n}\right)$$
$$\simeq n^{2(k-1)} \int_{0<x_1<\cdots<x_{k-1}<1} \frac{x_1^2 \cdots x_{k-1}^2}{(1-x_1)\cdots(1-x_{k-1})} (x_2-x_1)\cdots(1-x_{k-1}) \mathrm{d}x_1 \cdots \mathrm{d}x_{k-1}$$
(VIII.2.5)

$$\int_0^{1-\frac{1}{n}} \frac{\mathrm{d}x_1 \cdots \mathrm{d}x_{k-1}}{(1-x_1)\cdots(1-x_{k-1})} = \frac{1}{(k-1)!} \left( \int_0^{1-\frac{1}{n}} \frac{\mathrm{d}x}{1-x} \right)^{k-1} = \frac{(\ln(n))^{k-1}}{(k-1)!} \tag{VIII.2.6}$$

We can bound the numerator of (VIII.2.5), by bounding the product of square terms by $n^{-(2k-1)} \frac{(n-1)!^2}{(n-k)!^2}$. This arises from substituting the largest values possible for the $x_i$, and the difference terms by $\left(\frac{n-k-1}{n}\right)^{k-1}$, since this is the highest any difference

between two consecutive terms can be. This leaves us with:

$$F_n \leq \frac{1}{2}((n-1)!)^3 \sum_{k=1}^{n-2} \frac{\left(\frac{\ln(n)}{2n}\right)^{k-1}}{(k-1)!((n-k)!)^2} \tag{VIII.2.7}$$

ignoring the $(n-k)!$ in the denominator, the sum is bounded by an exponential, since each term is positive, and so we have

$$F_n \leq \frac{1}{2}((n-1)!)^3 n^{\frac{1}{2n}} \tag{VIII.2.8}$$

This makes the modulus of the terms in the generating series we want to understand the convergence of:

$$a_n = \frac{1}{2}\frac{n-1}{n}((n-1)!)^2 n^{\frac{1}{2n}} \tag{VIII.2.9}$$

This has zero radius of convergence. It makes it necessary to consider better approximations if it is convergent.

### VIII.2.4 Enumeration Problems

There is an alternative viewpoint about enumerating the flower graphs. The idea is that we can partition our set of vertices into the cycles or chains with the first cycle being labelled 1 and the subsequent chains being labelled in ascending order of the least element inside of them. Then we can connect the sets in the partition according to where the endpoints of the chains join to the main body of the graph. For example, the set labelled 2 has to have both ends joined to 1; the set 3 can have both to 1, both to 2 or one to both. We can understand the number of such structures by thinking inductively. When we add the $(n+1)$th set in our partition we can join it twice to any previous set or to any unordered pair of the previous $n$ sets. This gives us $n + \frac{n(n-1)}{2}$ possibilities, giving the inductive formula:

$$P_{n+1} = \left(n + \frac{n(n-1)}{2}\right) P_n = \frac{n(n+1)}{2} P_n \tag{VIII.2.10}$$

This then leads to:

$$P_n = \frac{n!(n-1)!}{2^n} \tag{VIII.2.11}$$

since $P_1 = 1$. The exponential generating function:

$$P(x) = \sum_{n=1}^{\infty} \frac{P_n}{n!} x^n \tag{VIII.2.12}$$

then has zero radius of convergence. Since the $P$ structure gives an overall structure for the flower graphs, it implies that the flower graph exponential generating function is divergent.

**Remark 42.** *It isn't entirely clear what the relationship between $F$ and $P$ is, since we have to make restrictions on the partitions in that the first set must have cardinality at least three. Also we need to understand the substructure, that is how the structure works on the chains and cycles. The problem arises from the fact that the set with* 1 *in it has to be different to the others. It may be a point to consider whether just giving the set with* 1 *in it the same structure as the rest would actually give smaller numbers for each term in the exponential generating series, thus meaning we can consider this structure to give zero radius of convergence for flower structures.*

## VIII.3 The Ree-Hoover Expansion

The Ree-Hoover expansion was introduced in the paper [RH64b] and has been successfully applied in order to obtain either analytically or computationally virial coefficients up to the tenth order for purely hardcore interactions [RH67, RH64a, ClMc06, ClMc05] in various dimensions. Furthermore, it has also been used to explore analytic properties of the virial expansion in [ClMc04]. A key advantage of the approach is the simplicity in the set up. It also reflects a lot of the combinatorial cancellations, which are indicated in Chapter III.

Recall the edge weight used in the Mayer expansion:

$$f_{i,j} = e^{-\beta\Phi(x_i - x_j)} - 1 \tag{VIII.3.1}$$

The reformulation of the virial expansion in [RH64b], introduces the tilde function, through the important relationship:

$$1 = \tilde{f_{i,j}} - f_{i,j} \tag{VIII.3.2}$$

This is equivalent to:

$$\tilde{f_{i,j}} := e^{-\beta\Phi(x_i - x_j)} \tag{VIII.3.3}$$

For each missing edge in a two-connected graph, the factor $1 = \tilde{f_{i,j}} - f_{i,j}$ is introduced, to develop new graphs, which include a dashed line to denote the $\tilde{f}$-edges.

The virial expansion can be cast as:

$$\beta P = \rho + \sum_{k=2}^{\infty} c_k \rho^k \tag{VIII.3.4}$$

where

$$c_k := \frac{(k-1)}{k!} \int V_k(\mathbf{x}_1, \cdots, \mathbf{x}_k)\, \mathrm{d}^D\mathbf{x}_1 \cdots \mathrm{d}^D\mathbf{x}_{k-1} \tag{VIII.3.5}$$

where the Husimi function $V_k$ is the sum of all labelled two-connected Mayer graphs with $k$ points. The standard approach is to group all labelled graphs according to their isomorphism type and define:

i) $S_i[n]$ - the integral of an unlabelled graph of (isomorphism) type $i$ and $n$ points

ii) $\sigma_i[n]$ - the number of labellings of a graph of type $i$ with $n$ points

This provides us with the interpretation:

$$c_n = -\frac{n-1}{n!} \sum_i \sigma_i[n] S_i[n] \tag{VIII.3.6}$$

All graphs when the replacement $1 = \tilde{f_{i,j}} - f_{i,j}$ is made are complete, with edges being either dashed $\tilde{f_{i,j}}$ edges or regular $f_{i,j}$ edges. The contribution of the diagram whose regular edges are the edges found in $S_i[n]$, is denoted $\tilde{S}_i[n]$. The star content of a diagram $\tau_i[n]$ is the combinatorial factor that is created from the fact that for every two-connected subgraph, we can form $\tilde{S}_i[n]$, by adding the appropriate $\tilde{f}$ and $f$-edges. Hence, we end up with the expansion:

$$c_n = -\frac{n-1}{n!} \sum_i \tau_i[n] \sigma_i[n] \tilde{S}_i[n] \tag{VIII.3.7}$$

The rule for counting the star content is:

*count the number of labelled two-connected graphs, which can be formed by successively removing an even number of edges from the Ree-Hoover diagram and subtract from this the number of labelled two-connected diagrams that are formed from removing an odd number of edges from the Ree-Hoover diagram. The resulting number is precisely* $\tau_i[n]$.

For the ten two-connected graphs on four vertices, the Ree-Hoover expansion reduces to four graphs, where the complete graph has star content $-2$. We note here that three of the graphs which are cancelled can be understood as maximal labelled graphs with **h**-vectors $(-1, 0, -1)$, $(-1, -1, 2)$ and $(1, 0, 0)$ in terms of the set up in Chapter III. This also follows through for the general case as noted below.

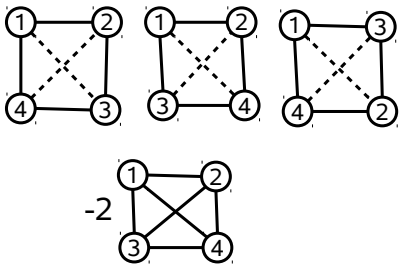



Figure VIII.3: The Contributing Diagrams for the Ree Hoover Expansion at $n = 4$

We use the notation $SC(G)$ to denote the star content of a specific graph $G$ and $\mathcal{B}_i(G)$ to denote the collection of two-connected graphs formed by removing $i$ edges from $G$. We can then write the star content of a graph $G$ as:

$$SC(G) = \sum_{i=1}^{|E|} (-1)^i |\mathcal{B}_i(G)| \tag{VIII.3.8}$$

We make the simple observations that $|\mathcal{B}_0(G)| = 1$ if and only if $G$ is two-connected and $|\mathcal{B}_i(G)| = 0 \ \forall i > 0$ if and only if $G$ is minimally two-connected.

We may interpret $\mathcal{B}_i$ as a mapping from sets of graphs to sets of graphs, where we define the extension:

$$\mathcal{B}_i(\{G_1, \cdots, G_n\}) := \bigcup_{j=1}^{n} \mathcal{B}_i(G_j) \tag{VIII.3.9}$$

In this case it is clear that $\mathcal{B}_i \circ \mathcal{B}_j = \mathcal{B}_{i+j}$. We also have the recursive formula for the size of the sets:

$$|\mathcal{B}_{i+1}(G)| = \frac{1}{i+1} \sum_{G' \in \mathcal{B}_1(G)} |\mathcal{B}_i(G')| \tag{VIII.3.10}$$

There is also a useful connection to Chapter III, regarding how to calculate the star content. We notice that for **h**-values, as given in Section III.3, which contain elements different from 0 or $-1$ or are non-decreasing sequence, the maximal labelled graph corresponding to such a **h**-vector will have zero star content as it is shown that the alternating sum of these graphs all cancel. This works for labelled graphs and can then be interpreted as reducing the number of graphs for a particular isomorphism type. Furthermore, it indicates that the star content of a complete graph is necessarily $-(n-2)!$.

The main application of the Ree-Hoover expansion is to obtain virial co-

efficients for hard-core models in $D$-dimensions. The advantage the Ree-Hoover expansion plays is that some graphs are eliminated via having zero star content and other graphs can be eliminated due to geometrical considerations, when the collection of $f_{i,j}$ and $\tilde{f_{i,j}}$-functions is incompatible for a particular diagram. The tilde functions enforce the two labels to be of distance greater than 1 apart, whereas the Mayer functions enforce the two labels to be less than 1 apart.

For graphs with the same tilde graph, but different number of vertices, we have the relationship for star content:

$$\tau_i[n] = (-1)^{\binom{n}{2}-\binom{m}{2}}\tau_j[m]\frac{(n-1)!}{(m-2)!} \tag{VIII.3.11}$$

Indeed, if we consider a Ree-Hoover graph and remove a vertex, then, we see that for the subgraph to have the same tilde subgraph, we must remove a vertex which has $f$-edges to each other vertex. Each contributing term in the sum for $n-1$ appears in the sum for $n$, but with the possibility of at least two edges being added from the 'removable' point to the graph. This will give the alternating sum of factors $\binom{n-1}{k}(-1)^{n-1-k}$ for $2 \leq k \leq n-1$. Of course the full sum is zero and so this truncated version is $(-1)^n + (-1)^{n-1}(n-1) = (-1)^{n-1}(n-2)$.

$$\tau_i[n] = (-1)^{n-1}(n-2)\tau_j[n-1] \tag{VIII.3.12}$$

If we let $m$ be the least number such that we have a Ree-Hoover graph with the same $\tilde{f}$ -subgraph, then we can formulate the virial coefficients as:

$$B_n = -\frac{1}{n}\sum_i (-1)^{\binom{n}{2}-\binom{m}{2}}\frac{\tau_j[m]\sigma_i[n]\tilde{S}_i[n]}{(m-1)!} \tag{VIII.3.13}$$

Note also that $\sigma_i[n] = \binom{n}{p}\sigma_j[m]$, where $p$ is the number of vertices connected by $\tilde{f}$ in the corresponding $\tilde{f}$-graph. We can therefore, when given an $\tilde{f}$-graphs, sum over $n$, the factor:

$$\frac{(-1)^{\binom{n}{2}}\tilde{S}_i[n](n-1)!}{(n-p)!} \tag{VIII.3.14}$$

and have an overall $n$-independent factor:

$$\frac{\tau_j[m](-1)^{1-\binom{m}{2}}}{(m-2)!p!} \tag{VIII.3.15}$$

Thus the summation can be separated into the separate sums for different tilde graphs. This approach does not seem to have been taken in the literature.

## Conclusions & Open Questions

In order to find a way to estimate the virial coefficients, this section has explored the avenue of using minimal two-connected graphs in order to emulate the Penrose tree-construction to achieve bounds. The literature has given particular starting points such as using the original equations governing the thermodynamic properties and making approximations here then interpreting them graphically, in the hope that one may understand the final expansion through the iterative procedure. The problems lie in the graphs which are discluded and require particular questions detailed in the papers [Ste63, LGB59].

The key route would be to allow accuracy to be retained as much as possible in evaluating the coefficients. This is reflected in the success of the Ree-Hoover reformulation of the virial equation of state, which uses a clever trick of setting $1 = \tilde{f_{i,j}} - f_{i,j}$ in order to reduce the number of graphs considered and to furthermore obtains graphs which can easily be neglected in the case of hardcore gases. The work done in Chapter III can be used to understand the cancellation of certain coefficients as this reformulation relies upon elements of the Tonks gas.

This chapter has also presented the issues around using the minimal two-connected graphs and obtaining suitable partitions. Partitions of the form of Penrose, where we are able to express everything in terms of a minimal and maximal graph are not possible in this case. The number of these minimal graphs also appears to be divergent and so making the approximation would not be as profitable as in the case of connected graphs. Further cancellations must appear somewhere else in the virial expansion and computing these should focus rather on the ideas of Chapter III. Ree-Hoover is noted not to deal with every cancellation in the virial expansion and it would be helpful to devise a method for doing this.

Further open questions of interest about the virial expansion include whether it is possible to find negative $c_k$ for hard discs and spheres, which would have consequences for the interpretation for the point at which the virial expansion fails to converge. It would be beneficial to compute the fraction of Ree-Hoover diagrams that vanish identically for large $k$. An analytic expression for $c_k$ for $k \geq 5$ has not yet been attained. Improved accurate reformulations may aid in doing so, or else we can improve approximations.

# Chapter IX

# Multispecies Virial Expansion

A further development to the work already presented in this thesis is to understand the generalisation to many different types of particle. The result of this is to have different fugacity parameters $(z_i)_{i\in I}$ and densities $(\rho_i)_{i\in I}$, where $I$ represents a label set for the particle type. The work presented in this chapter was mostly done in collaboration with Jansen Tsagkarogiannis and Ueltschi and is found in the paper [JTTU14].

The motivation for considering the multispecies generalisation is that they are a very natural object as mixtures in Chemistry. They also arise as effective models, when considering certain phenomena related to the order-disorder transition in alloys [Fuc42]. Furthermore, the notions of a multicomponent system appear in the treatment of a monatomic gas as a mixture of droplets (groups of particles close in space) [Hill56]. These droplets may comprise of arbitrarily many droplets and occur with arbitrary size, which motivates the need for an arbitrary number of species.

Simultaneous to the study of monatomic gases, detailed in Chapters I and II, there was a serious investigation of multi-component systems, which can, for example, be found in [May39] for the case of a two-component system and later in [Fuc42] for a mixture with an arbitrary (but finite) number of different components. Although briefly mentioned already in [Fuc42], the complete study of the convergence of the activity expansion comes later in [BaLe64] for mixtures of finitely many components. Further background material can be found in [May37, MA37, HaMa38, BoFu38, KaUh38].

The generalisation of virial expansions and approximate equations of state from the monatomic gases to mixtures has important intrinsic difficulties. In the paper [HeLe70], the van der Waals equation for binary mixtures of hard spheres

comes in different versions, distinguished by different 'mixing rules' for obtaining effective parameters of the mixture. Such problems are also noted in [LeRo64] for the Percus-Yevick and virial equations of state. Section IX.1.presents how a naïve generalisation of the van der Waals equation of state encounters particular difficulties.

In [Fuc42], it is indicated that even solving the case for binary mixtures gives no indication on how to deduce results for more than two components. In particular, the reduction of the cluster integrals to the irreducible cluster integrals is not clear even in this case. The approach of this paper and the context of the other background papers is presented in Section IX.2. This gives the key ideas that are useful when considering the case of infinitely many particle types. Furthermore, this section indicates the various applications found for a multispecies virial expansion.

From finitely many particle types, it would be interesting to generalise to infinitely many types of particles. A useful result would be an estimate on the domain of convergence.

Section IX.4 presents the virial expansion for countably many types of particle, addressing the result for convergence. The existence of a pressure function that depends on countably many fugacity parameters is assumed. The pressure function yields countably many densities (that are also functions of fugacities) and the goal is to write the pressure as function of densities. In the case of one parameter, this can be done using Lagrange inversion [LePen64, Rue69].

The analytic method of Lagrange inversion is also a standard tool in combinatorics [BLL98]. In the case of multicomponent systems, the key ingredient is Good's generalisation of the Lagrange inversion to several variables [Good60]. The Lagrange-Good inversion has attracted attention in a variety of contexts [Abd03a, Bru83, Ges87, EhMe94]. Faris has recently noticed its relevance for the virial expansion [Far12b]. Our main focus is on the convergence of the expansion. This is covered in Section IX.3. The main result is the existence of a non-trivial domain of convergence for the expansion of the pressure as function of densities. A novel feature is to use Lagrange-Good for proving convergence.

The work of Mayer [MMay40] initiated the connection between cluster and virial expansions with connected (Section I.2) and two-connected graphs (Section II.2) respectively. Section IX.5 deals with a gas of classical particles with two-body interactions. There are many species of particles. Under some given assumptions, the virial coefficients are given in terms of two-connected graphs (irreducible cluster integrals). This is done through the dissymmetry theorem for coloured weighted graphs. This result holds whenever the weights satisfy a block factorisation property.

Early derivations for systems with one, two, or finitely many components can be found in [BoFu38, May39, Fuc42]. This connection has been further developed in the work of Leroux and collaborators [DLL07], leading to modern proofs for the expression of virial coefficients. This was generalised by Faris to the case of many species [Far12b]. This chapter formulates sufficient conditions on the interactions that guarantee the convergence of the virial expansion.

The results have relevance beyond statistical mechanics. Indeed, the main result can be formulated as an inverse function theorem for functions between Banach spaces that are not necessarily Fréchet-differentiable (Section IX.7). In addition, the result can be applied in the original context of the Lagrange-Good inversion formula: Good motivated his work by stochastic branching processes and combinatorics of coloured trees [Good60, Good65]. Recursive properties of trees lead to functional equations between generating functions. When combined with the inversion formula they yield expressions of probabilities or tree cardinalities as contour integrals; Good explicitly computed some of those integrals. The result yields bounds for cardinalities without having to compute the integrals, which is useful when explicit computations prove too complicated.

In Section IX.6, a mixture of rigid molecules is considered. Using the results on the convergence of the cluster expansion given in [Uel04, PoUe09], the mixture is shown to meet the conditions of Section IX.4. It is also important to emphasise that generalising this to more complicated systems provided inherent difficulties to bound particular integrals of complex valued functions along a contour.

## IX.1 A Generalisation of the van der Waals Equation of State

In [HeLe70], it is emphasised that it is difficult to measure experimentally the interaction of unlike molecules in order to obtain an equation of state for a mixture. Instead a notion of the depth of the interaction between unlike particles is introduced and is denoted by $\varepsilon_{1,2}$. The key issue is that the behaviour of a virial expansion through the van der Waals equation is sensitive to small changes in this parameter. The principle of corresponding states is frequently used to describe such mixtures, by considering an average interaction. The key problem that is presented is there is a possibility of many different mixing rules to obtain the $\rho_1\rho_2$ coefficient depending on how the interaction is determined. However, in the work of Jansen and Tsagkarogiannis [JaTs13], a positive result is given for the many species Tonks gas, where the virial expansion is precisely the natural generalisation of the van der Waals equation

of state.

The naïve generalisation to multispecies for the van der Waals equation of state is:

$$\left(P + \frac{N_1^2 a_{11} + 2N_1 N_2 a_{12} + N_2^2 a_{22}}{V^2}\right)(V - b_1 N_1 - b_2 N_2 + \mathcal{O}(N\rho)) = (N_1 + N_2)kT \tag{IX.1.1}$$

The first term corresponds to the weak potential/high-temperature effect of the potentials between either particles of the same type or particles of two different types. The $N_1 N_2$ is the number of pairwise interactions. The idea of a volume exclusion effect or the use of hardcore conditions is less transparent in this situation. The volume exclusion effect is difficult to manage or understand and we consider a linear approximation to this, since, as we will see, any effect of higher order would not effect the second virial coefficient on which we now focus. We rearrange (IX.1.1) to obtain:

$$(P + \rho_1^2 a_{11} + 2\rho_1\rho_2 a_{12} + \rho_2^2 a_{22})(1 - b_1\rho_1 - b_2\rho_2 + \mathcal{O}(\rho^2)) = (\rho_1 + \rho_2)kT \tag{IX.1.2}$$

We can then make the appropriate algebraic manipulations:

$$\begin{aligned} \beta P &= (\rho_1 + \rho_2)(1 + b_1\rho_1 + b_2\rho_2 + \mathcal{O}(\rho^2)) = \beta\rho_1^2 a_{11} - 2\beta\rho_1\rho_2 a_{12} - \beta\rho_2^2 a_{22} \\ \beta P &= \rho_1 + \rho_2 + \rho_1^2(b_1 - \beta a_{11}) + \rho_2^2(b_2 - \beta a_{22}) + \rho_1\rho_2(b_1 + b_2 - 2\beta a_{12})\mathcal{O}(\rho^3) \end{aligned} \tag{IX.1.3}$$

We can draw the same comparison for the multispecies van der Waals as was made for the single species in Section II.1, since the only difference is the different potentials we may integrate over. This derivation indicates that the hardcore conditions for type 1 and type 2 would have to sum to give that of the interaction between 1 and 2, but this will not always be true. We can consider the case of having the two types being protons and electrons, which repel like particles at short range, but are attractive for unlike particles. The van der Waals equation of state thus does not have a straightforward universal application to mixtures.

## IX.2 Background Material for Multi-species Thermodynamics

They key background material to the multi-species generalisation presented in this chapter is the work of Fuchs [Fuc42], in which Mayer's formulation of the cluster and virial expansions is extended to finitely many species case. Indeed, to obtain the interpretation of two-connected graphs for the virial expansion, the concept

of an incomplete irreducible cluster is introduced, where the irreducible cluster is incomplete due to the fact that one articulation point is left out and this has a particular type, which marks the 'type' of the irreducible cluster. This is analogous to the method of proof for the dissymmetry theorem presented in Section IX.5, in which an articulation point is left out that is closest to the bc-centre of the tree, which is the canonical approach in combinatorial species. Fuchs leaves an arbitrary articulation point undetermined and then continues from there. A step by step interpretation of the combinatorics of forming connected graphs from these irreducible components is given, which reflects the Lagrange-Good inversion given in Section IX.3:

$$l_1 b(\mathbf{l}) = \sum_{(\mu_i(\boldsymbol{\nu}))} \left| \delta_{i,j} - \frac{1}{l_j} \sum_{(\nu)} (\nu_i - \delta_{i,j}) \mu_j(\boldsymbol{\nu}) \right|_A \prod_{i=1}^{A} \prod_{(\nu)} \frac{(l_i \nu_i B(\boldsymbol{\nu}))^{\mu_i(\boldsymbol{\nu})}}{\mu_i(\boldsymbol{\nu})!} \tag{IX.2.1}$$

where $b(\mathbf{l})$ are the cluster coefficients and $B(\boldsymbol{\nu})$ are the virial coefficients. The vector $\mathbf{l}$ denotes the coefficient of

$$\prod_{i=1}^{A} z_i^{l_i} =: \mathbf{z}^{\mathbf{l}} \tag{IX.2.2}$$

in the cluster expansion and similarly for $B(\boldsymbol{\nu})$ in the virial expansion for $\boldsymbol{\rho}$. $A$ is the number of species. The modulus indicates taking the determinant of the matrix with the defined $(i,j)$-entries. The summation is over all collections of integers $\mu_i(\boldsymbol{\nu}) \geq 0$ such that:

$$\sum_{j=1}^{A} \sum_{\boldsymbol{\nu}} (\nu_i - \delta_{i,j}) \mu_j(\boldsymbol{\nu}) = \begin{cases} l_i & \text{if } i \neq 1 \\ l_i - 1 & \text{if } i = 1 \end{cases} \tag{IX.2.3}$$

The above is also written in terms of generating functions through integral representations of the appropriate coefficients and the formula may also be obtained via contour integrals of complex analysis.

Another key issue that is emphasised in this paper is the notion of associated radii of convergence, first introduced by Borel [Bor01], for multidimensional expansions. Constants $K_1 \cdots K_{A-1}$ are introduced and the ratios of the $N_i$-number of particles of species $i$ are modified in order to give the largest limit of

$$\left( Q_\tau(\mathbf{N}) \prod_{i=1}^{A-1} K_i^{N_i} \right)^{\frac{1}{N}}$$

as $\sum_{i=1}^{A} N_i = N \to \infty$. This is denoted $\lambda(\mathbf{K})$. $Q_\tau(\mathbf{N})$ is the canonical partition function and the coefficient of $\mathbf{z}^{\mathbf{N}}$ in the cluster expansion. The set of associated radii of convergence are $\frac{X_i}{K_i} = X_A = \frac{1}{\lambda(K)}$. These replace the single radius of convergence in the case of one dimension.

The paper of Lebowitz and Rowlinson [LeRo64] generalises the Percus-Yevick equation and focuses on the case of the hard sphere. The paper of Baert and Lebowitz [BaLe64] generalises the cluster expansion bounds of Groeneveld and Penrose to the (finite) multispecies expansion, giving the requirement:

$$\sum_{i=1}^{A} |z_i| \leq \left(\exp\left(1 + 2\beta B\right) C_m(\beta)\right)^{-1} \tag{IX.2.4}$$

where $C_m(\beta)$ is the maximal temperedness:

$$C_m(\beta) = \max_{(i,j)} \int |\exp(-\beta \Phi_{i,j}(\mathbf{r})) - 1| \, \mathrm{d}\mathbf{r} \tag{IX.2.5}$$

with $\Phi_{i,j}(\mathbf{r})$ the pair potential between a particle of species $i$ and $j$.

$B$ is the universal stability bound, such that for any $s \in \mathbb{N}$, collection of particle types, denoted by $c : [s] \to [A]$ and locations $(\mathbf{x}_i)_{i \in [s]}$, we have:

$$\sum_{1 \leq i < j \leq s} \Phi_{c(i),c(j)}(|\mathbf{x}_i - \mathbf{x}_j|) \geq -sB \tag{IX.2.6}$$

Furthermore, multispecies has been a useful tool in the study of Anyon models by Isakov Mashkevich and Ouvry [Isa94, IMO95, IsMa97], in which a similar derivation of virial coefficients is made to low order and through algebraic relationships.

## IX.3 Lagrange-Good Inversion

The history of Lagrange inversion and attempts to extend it to many dimensions includes many different techniques and approaches. It is a very rich subject, which is full of various connections within Mathematics.

The first main proof of a general finite dimensional Lagrange inversion is given by Good [Good60, Good65], when considering the probability generating function of the sizes of generations in a tree consisting of $C$ distinct species or colours. The theorem of Good is given below:

**Theorem IX.3.1** (Lagrange-Good Inversion)**.** *Let*

$$\mathbf{z} = \boldsymbol{\zeta} \wedge \boldsymbol{f}(\mathbf{z}) \tag{IX.3.1}$$

*each component of the vectors indicating a separate equation and the* $\wedge$ *operation meaning to multiply vectors componentwise. Let* $\Psi(\boldsymbol{z}, \boldsymbol{\zeta})$ *be a meromorphic function in a neighbourhood of* $\mathbf{z} = \boldsymbol{\zeta} = \mathbf{0}$*. Then:*

$$[\boldsymbol{\zeta}^{\mathbf{m}}]\Psi(\mathbf{z}(\boldsymbol{\zeta}), \boldsymbol{\zeta}) = [\mathbf{z}^{\mathbf{m}}] \left( \Psi(\mathbf{z}, \boldsymbol{\zeta}(\mathbf{z}))\mathbf{f}(\mathbf{z})^{\mathbf{m}} \left| \delta_{i,j} - \frac{z_i}{f_i(\mathbf{z})} \frac{\partial f_i(\mathbf{z})}{\partial z_j} \right|_{i,j \in [C]} \right) \tag{IX.3.2}$$

There is another useful paper of Merlini Sprugnoli and Verri [MSV06], which indicates applications of the Lagrange inversion formula and other interpretations of Lagrange inversion, including computing generating functions of sequences; performing coefficient extraction of a formal power series; comparing combinatorial sums; and performing inversion of identities. Their formulation for the one-dimensional case is stated below due to the simplicity of applying it.

**Theorem IX.3.2** (Merlini Sprugnoli Verri formulation of Lagrange inversion)**.** *Let* $w = w(t)$ *be defined implicitly by* $w = t\varphi(w)$*, where* $\varphi(t)$ *is a formal power series with* $\varphi(0) \neq 0$*, then we have:*

$$[t^n]w(t)^k = \frac{k}{n}[t^{n-k}]\varphi(t)^n \tag{IX.3.3}$$

This is often called the Bürmann inversion formula.

In combinatorics there is also the interpretation of formal power series as matrices and as Riordan arrays. A Riordan array comes from a pair of formal power series $(d(t), h(t))$, where $d(0) \neq 0$ and $h(0) - 0$ with $h'(0) \neq 0$. This generates a lower triangular array $d_{n,k} := [t^n]d(t)h(t)^k$. If $(f_n)_{n \in \mathbb{N}}$ is a sequence of numbers corresponding to the generating function $f(t)$, we have:

$$\sum_{k=0}^{n} d_{n,k} f_k = [t^n]d(t)f(h(t)) \tag{IX.3.4}$$

A species interpretation is also given for the multispecies Lagrange inversion, which extends that presented in Section VI.2, in [GL95] for finitely many species and also in [EhMe94] for an arbitrary but countable number of species. The key feature is the definition of composition in coloured species case and its relationship with plethysm and the umbral calculus.

Another perspective is given by Abdesselam in [Abd03a], giving an inter-

pretation through complex bosonic field theory, involving transformation of fields: $\bar{\phi} \to \bar{\phi}$ and $\phi \to V(\phi) = \phi - xg(\phi)$. Another paper [Abd03b] links this consistent interpretation with the related Jacobian conjecture.

An operator interpretation of achieving Lagrange inversion is presented in [Kra88], giving orthogonality relations through properties of sequences $(f_k)_{k\in\mathbb{Z}}$ and $(\tilde{f}_k)_{k\in\mathbb{Z}}$ with $(f_k, \tilde{f}_n) = \delta_{n,k}$ with the inner product defined by:

$$[z^0]\left(f^k(z)\left(\frac{zf'(z)}{f^{n+1}(z)}\right)\right) = \delta_{n,k} \tag{IX.3.5}$$

In [EhMe94], the infinite version of Lagrange-Good inversion is given as:

**Theorem IX.3.3** (Infinitely Variated Good's Inversion Formula)**.** *Let* $\mathbf{f}(\mathbf{x})$ *be a* ***summable*** *collection of formal power series and let* $\mathbf{G}(\mathbf{x})$ *be a collection of formal power series, such that for all* $i \in \mathbb{N}$

$$f_i(\mathbf{x}) = x_i(G_i \circ \mathbf{f}(\mathbf{x})) \tag{IX.3.6}$$

*Let* $J = \{i \in \mathbb{N} | n_i \neq 0\}$ *and assume* $\mathbf{n} \geq \mathbf{k}$ *Then we have that:*

$$[\mathbf{x}^{\mathbf{n}}]\prod_{i\in\mathbb{N}} f_i(\mathbf{x})^{k_i} = [\mathbf{x}^{\mathbf{n}-\mathbf{k}}]\left|\delta_{i,j}G_i(\mathbf{x})^{n_i} - x_j\frac{\partial G_i(\mathbf{X})}{\partial x_j}G_i(\mathbf{x})^{n_i-1}\right|_{i,j\in J} \tag{IX.3.7}$$

This formula can be recast in a version amenable to the context for virial and cluster expansions. The key relationship is:

$$\rho_i = z_i \times \frac{\partial P}{\partial z_i} \tag{IX.3.8}$$

The thermodynamic variables are related to the functions in (IX.3.6) by making the identifications:

$$f_i = z_i \quad x_i = \rho_i \quad \text{and} \quad G_i = \frac{1}{\frac{\partial P}{\partial z_i}}$$

We may recast (IX.3.7) by substituting $\rho_i$ for $x_i$ and $z_i$ for $f_i$ on the left hand side, but $z_i$ for $x_i$ and of course $\frac{1}{\frac{\partial P}{\partial z_i}}$ for $G_i$ on the right hand side. This leads to the corollary:

**Corollary IX.3.4** (Thermodynamic version of Good's Inversion Formula)**.** *For activity parameters* $z_i$*, pressure* $P$ *and densities* $\rho_i$*, we have the following inversion*

*formula:*

$$[\boldsymbol{\rho}^{\mathbf{n}}]\mathbf{z}^{\mathbf{k}} = [\mathbf{z}^{\mathbf{n}-\mathbf{k}}] \left| \delta_{i,j} \left( \frac{1}{\frac{\partial P}{\partial z_i}} \right)^{n_i} - z_j \frac{\partial}{\partial z_j} \left( \frac{1}{\frac{\partial P}{\partial z_i}} \right) \left( \frac{1}{\frac{\partial P}{\partial z_i}} \right)^{n_i - 1} \right|_{i,j \in \mathbb{N}} \tag{IX.3.9}$$

## IX.4 General virial expansions

The paper [JTTU14] gives conditions for the convergence of the virial expansion, dependent on conditions on mutlispecies cluster expansion coefficients. It also indicates, in the context of [PoUe09], the example of rigid polymers as an application of these bounds. It draws upon connections with the coloured dissymmetry theorem and requirements for the multispecies virial coefficients to be written in terms of weighted coloured two-connected graphs analogous to the one species case.

### IX.4.1 Setting & results

Let $\boldsymbol{z} = (z_1, z_2, \dots)$ denote a sequence of complex numbers, where $z_i$ is interpreted as the fugacity of species $i$. Consider the formal series

$$p(\boldsymbol{z}) = \sum_{\boldsymbol{n}} b(\boldsymbol{n}) \boldsymbol{z}^{\boldsymbol{n}}, \tag{IX.4.1}$$

where the sum is over all multi-indices $\boldsymbol{n} = (n_1, n_2, \dots)$, $n_i \in \mathbb{N}$, with finitely-many non-zero entries and $\sum_{i \in \mathbb{N}} z_i \geq 1$. We assume that for all $i \in \mathbb{N}$, the coefficient $b(0, \dots, 1, 0, \dots)$ of $z_i$ in $p(\boldsymbol{z})$ is non-zero; this is the only condition needed to apply Lagrange inversion. In Section IX.6, we make the additional assumption that the coefficients are normalised to $b(0, \dots, 1, 0, \dots) = 1$, as is the case in most applications in statistical mechanics.

$p(\boldsymbol{z})$ is the pressure of the system with many species. A physically relevant quantity is the density $\rho_i$ of the species $i$, whose definition is

$$\rho_i(\boldsymbol{z}) = z_i \frac{\partial p}{\partial z_i}(\boldsymbol{z}). \tag{IX.4.2}$$

We do not suppose yet that the series for $p(\boldsymbol{z})$ is convergent, and the equation above should be understood in the sense of formal series. To be precise, $\rho_i(\boldsymbol{z})$ is the formal series with coefficients $n_i b(\boldsymbol{n})$.

The virial coefficients $c(\boldsymbol{n})$ are given by:

$$p(\boldsymbol{z}) = \sum_{\boldsymbol{n}} c(\boldsymbol{n}) \boldsymbol{\rho}(\boldsymbol{z})^{\boldsymbol{n}}. \tag{IX.4.3}$$

The coefficients $c(\boldsymbol{n})$ are well-defined in the sense of formal series. Observe that $[\boldsymbol{z}^{\boldsymbol{k}}]\boldsymbol{\rho}(\boldsymbol{z})^{\boldsymbol{n}} \neq 0$ only if $\boldsymbol{n} \leq \boldsymbol{k}$, i.e., if $n_i \leq k_i$ for all $i$. Then $(\boldsymbol{\rho}(\boldsymbol{z})^{\boldsymbol{n}})_{\boldsymbol{n}}$ is a "summable family" of formal series in the sense of Def. 2.2 in [EhMe94], and the coefficients of $\boldsymbol{z}^{\boldsymbol{k}}$ in (IX.4.3) satisfy

$$b(\boldsymbol{k}) = [\boldsymbol{z}^{\boldsymbol{k}}] \sum_{\boldsymbol{n}} c(\boldsymbol{n}) \boldsymbol{\rho}(\boldsymbol{z})^{\boldsymbol{n}} = \sum_{\boldsymbol{n} \leq \boldsymbol{k}} c(\boldsymbol{n}) [\boldsymbol{z}^{\boldsymbol{k}}] \boldsymbol{\rho}(\boldsymbol{z})^{\boldsymbol{n}}, \tag{IX.4.4}$$

the latter sum being finite. Then Eq. (IX.4.4) can be inverted recursively, and $c(k_1, \dots, k_\ell, 0, \dots)$ can be expressed in terms of $b(\boldsymbol{n})$ and $c(n_1, \dots, n_\ell, 0, \dots)$ with $n_i \leq k_i$, $i = 1, \dots, \ell - 1$, and $n_\ell < k_\ell$.

The goal is to control the convergence of the virial expansion, assuming convergence of the series $p(\boldsymbol{z})$. In the following, the statement "$|\log f(\boldsymbol{z})| \leq a$" means that there is a $A(\boldsymbol{z}) \in \mathbb{C}$ such that $f(\boldsymbol{z}) = \exp(A(\boldsymbol{z}))$ and $|A(\boldsymbol{z})| \leq a$, i.e., the precise choice of the branch of the logarithm is irrelevant.

**Theorem IX.4.1.** *Assume that there exist $0 < r_i < R_i$ and $a_i \geq 0$, $i = 1, 2, \dots$, such that*

- *$p(\boldsymbol{z})$ converges absolutely in the polydisc $D = \{\boldsymbol{z} \in \mathbb{C}^{\mathbb{N}} \mid \forall i \in \mathbb{N} : |z_i| < R_i\}$.*
- *$\left|\log \dfrac{\partial p}{\partial z_i}(\boldsymbol{z})\right| < a_i$ for all $i \geq 1$ and all $\boldsymbol{z} \in D$.*
- *$\displaystyle\sum_{i \geq 1} \sqrt{\frac{r_i}{R_i}} < \infty$ and $\displaystyle\sum_{i \geq 1} \frac{r_i a_i^2}{R_i} < \infty$.*

*Then there exists a constant $C < \infty$ (which depends on the $r_i$, $R_i$, $a_i$, but not on $\boldsymbol{n}$) such that*

$$|c(\boldsymbol{n})| \leq C \sup_{\boldsymbol{z} \in D} |p(\boldsymbol{z})| \prod_{i \geq 1} \Big( \frac{\mathrm{e}^{a_i}}{r_i} \Big)^{n_i}. \tag{IX.4.5}$$

The estimate for $c(\boldsymbol{n})$ guarantees convergence of the series $\sum_{\boldsymbol{n}} c(\boldsymbol{n}) \boldsymbol{\rho}^{\boldsymbol{n}}$ for all $\boldsymbol{\rho}$ in a polydisc

$$D' = \Big\{ \boldsymbol{\rho} \in \mathbb{C}^{\mathbb{N}} \mid \forall i \in \mathbb{N} : |\rho_i| < r_i \, \mathrm{e}^{-a_i} \, , \, \sum_{i \in \mathbb{N}} |\rho_i| \frac{\mathrm{e}^{a_i}}{r_i} < \infty \Big\}. \tag{IX.4.6}$$

We can also address the following question: Consider the functions $z_i(\boldsymbol{\rho})$ obtained by inverting (IX.4.2); for given $\boldsymbol{\rho} \in D'$, does $\boldsymbol{z}(\boldsymbol{\rho})$ belong to $D$, so that $p(\boldsymbol{z}(\boldsymbol{\rho}))$ is given by an absolutely convergent series? The following result provides a partial answer, as it guarantees convergence when $\boldsymbol{\rho}$ belongs to a smaller domain.

**Theorem IX.4.2.** *Under the same assumptions as in Theorem IX.4.1, we have*

$$\Big|[\boldsymbol{\rho}^{\boldsymbol{n}}]\frac{\boldsymbol{z}(\boldsymbol{\rho})^{\boldsymbol{k}}}{\boldsymbol{\rho}^{\boldsymbol{k}}}\Big| < C\prod_{i\geq 1}\frac{\mathrm{e}^{a_i(n_i+k_i)}}{r_i^{n_i}}.$$

The constant $C$ is the same as in Theorem IX.4.1, and the proof is similar. It can be found at the end of Section IX.4.2.

Let $i \in \mathbb{N}$ and choose $\boldsymbol{k} = (k_j) = (\delta_{i,j})$ in Theorem IX.4.2. We get that for all $\boldsymbol{\rho} \in D'$,

$$|z_i(\boldsymbol{\rho})| \leq C|\rho_i|\,\mathrm{e}^{a_i}\prod_{j\geq 1}\Big(1-\frac{\mathrm{e}^{a_j}\,|\rho_j|}{r_j}\Big)^{-1}, \tag{IX.4.7}$$

so that $\boldsymbol{z}(\boldsymbol{\rho}) \in D$ for $\boldsymbol{\rho}$ small enough. The inequality (IX.4.7) is the analogue of the bound $|\rho_k(\boldsymbol{z})| \leq |z_k|\exp(a_k)$, valid for $\boldsymbol{z} \in D$ under the assumptions of Theorem IX.4.1.

### IX.4.2 Lagrange-Good inversion & bounds of virial coefficients

The Lagrange-Good inversion formula IX.3.9 gives explicit expressions for $c(\boldsymbol{n})$ and $[\boldsymbol{\rho}^{\boldsymbol{n}}]\boldsymbol{z}^{\boldsymbol{k}}$, which can then be estimated. It gives the algebraic relationship between the coefficients of the related power series and the intention is to use this to come up with an analytic understanding in terms of bounds.

Let $N$ be the largest index $i$ such that $n_i \neq 0$, and consider the $N\times N$ matrix

$$M(\boldsymbol{z}) = \left(\delta_{ij} + \frac{z_i\frac{\partial^2 p}{\partial z_i\partial z_j}}{\frac{\partial p}{\partial z_i}}\right)_{1\leq i,j\leq N}. \tag{IX.4.8}$$

We use Eq. (4.5) of [Ges87] to get

$$[\boldsymbol{\rho}^{\boldsymbol{n}}]\Phi(\boldsymbol{z}(\boldsymbol{\rho})) = [\boldsymbol{z}^{\boldsymbol{n}}]\Phi(\boldsymbol{z})\frac{1}{\big(\frac{\partial p}{\partial \boldsymbol{z}}\big)^{\boldsymbol{n}}}\det M(\boldsymbol{z}). \tag{IX.4.9}$$

Here, we used the notation $\big(\frac{\partial p}{\partial \boldsymbol{z}}\big)^{\boldsymbol{n}} = \prod_i\big(\frac{\partial p}{\partial z_i}\big)^{n_i}$. We employ the formula above with $\Phi(\boldsymbol{z}) = p(\boldsymbol{z})$ and $\Phi(\boldsymbol{z}) = \boldsymbol{z}^{\boldsymbol{k}}/\boldsymbol{\rho}(\boldsymbol{z})^{\boldsymbol{k}}$. In [Ges87], the formula has been proved for finitely many species; a proof for infinitely many species is given in [EhMe94] (Theorem 4). Note, however, that we can apply the finitely many species version in our context because we only need it for $\boldsymbol{n}$ with finitely many non-zero entries.

In order to estimate the coefficients of the right side, we use Cauchy's formula and get upper bounds on the various terms. We start with the determinant in (IX.4.9).

**Lemma IX.4.3.** *Under the assumptions of Theorem IX.4.1, there exists a constant $C < \infty$ such that for all $\boldsymbol{z}$ with $|z_i| = r_i$, and all $N$, we have*

$$|\det M(\boldsymbol{z})| \leq C.$$

*Proof.* We start by expanding $\det M(\boldsymbol{z})$ in terms of determinants of minors. We denote by $\mathcal{S}[N]$ and $\mathcal{S}[J]$ denoting the set of permutations on $[N]$ and on $J \subset [N]$, respectively. The expansion is:

$$\begin{aligned}
\det M(\boldsymbol{z}) &= \sum_{\sigma \in \mathcal{S}[N]} \operatorname{sgn}(\sigma) \prod_{i=1}^{N} \left( \delta_{i,\sigma(i)} + \frac{z_i \frac{\partial^2 p}{\partial z_i \partial z_{\sigma(i)}}}{\frac{\partial p}{\partial z_i}} \right) \\
&= \sum_{\sigma \in \mathcal{S}[N]} \operatorname{sgn}(\sigma) \sum_{J \subset [N]} \prod_{i \in J^c} \delta_{i,\sigma(i)} \prod_{i \in J} \frac{z_i \frac{\partial^2 p}{\partial z_i \partial z_{\sigma(i)}}}{\frac{\partial p}{\partial z_i}} \\
&= \sum_{J \subset [N]} \sum_{\sigma \in \mathcal{S}[J]} \operatorname{sgn}(\sigma) \prod_{i \in J} \frac{z_i \frac{\partial^2 p}{\partial z_i \partial z_{\sigma(i)}}}{\frac{\partial p}{\partial z_i}} \\
&= \sum_{J \subset [N]} \left| \frac{z_i \frac{\partial^2 p}{\partial z_i \partial z_j}}{\frac{\partial p}{\partial z_i}} \right|_{i,j \in J},
\end{aligned} \tag{IX.4.10}$$

where the summand corresponding to $J = \emptyset$ is by definition equal to 1. Let $u_i$ be non-zero numbers to be determined later. We use the identity $\det M = \det DMD^{-1}$ with $D$ the diagonal matrix with entries $u_i^{-1} \frac{\partial p}{\partial z_i}$ in the diagonal, and we get

$$\det M(\boldsymbol{z}) = \sum_{J \subset [N]} \left| z_i \frac{u_j}{u_i} \frac{\partial}{\partial z_i} \log \frac{\partial p}{\partial z_j} \right|_{i,j \in J}. \tag{IX.4.11}$$

From Hadamard's inequality, we get the upper bound

$$|\det M(\boldsymbol{z})| \leq \sum_{J \subset [N]} \prod_{i \in J} \left( \sum_{j \in J} |z_i|^2 \frac{u_j^2}{u_i^2} \left| \frac{\partial}{\partial z_i} \log \frac{\partial p}{\partial z_j} \right|^2 \right)^{1/2}. \tag{IX.4.12}$$

By Cauchy's formula, choosing the contour around $z_i$ with radius $R_i - r_i$, and using the bound on the logarithm of $\frac{\partial p}{\partial z_j}$, we get

$$\left| \frac{\partial}{\partial z_i} \log \frac{\partial p}{\partial z_j}(\boldsymbol{z}) \right| = \left| \frac{1}{2\pi \mathrm{i}} \oint \frac{\log \frac{\partial p}{\partial z_j}(\hat{\boldsymbol{z}}^{(i)}, w)}{(w - z_i)^2} \mathrm{d}w \right| \leq \frac{a_j}{R_i - r_i}, \tag{IX.4.13}$$

where $\hat{\boldsymbol{z}}^{(i)}$ is the vector $\boldsymbol{z}$ without the $z_i$ term. Then

$$\begin{aligned} |\det M(\boldsymbol{z})| &\le \sum_{J\subset[N]} \prod_{i\in J} \Big[ \frac{r_i}{u_i(R_i-r_i)} \Big( \sum_{j\in J} u_j^2 a_j^2 \Big)^{1/2} \Big] \\ &\le \prod_{i=1}^{N} \Big[ 1 + \frac{r_i}{u_i(R_i-r_i)} \Big( \sum_{j\ge 1} u_j^2 a_j^2 \Big)^{1/2} \Big] \\ &\le \exp\Big[ \sum_{i\ge 1} \frac{r_i}{u_i(R_i-r_i)} \Big( \sum_{j\ge 1} u_j^2 a_j^2 \Big)^{1/2} \Big]. \end{aligned} \tag{IX.4.14}$$

Choosing $u_j = \sqrt{r_j/R_j}$, the expression above is finite by the assumptions of the lemma. □

*Proof of Theorem IX.4.1.* We use the Lagrange-Good inversion (IX.4.9) and Cauchy's formula and we write

$$c(\boldsymbol{n}) = \Big( \prod_{i=1}^{N} \frac{1}{2\pi \mathrm{i}} \oint \frac{\mathrm{d}z_i}{z_i^{n_i+1}} \Big) p(\boldsymbol{z}) \frac{1}{\big(\frac{\partial p}{\partial \boldsymbol{z}}(\boldsymbol{z})\big)^{\boldsymbol{n}}} \det M(\boldsymbol{z}), \tag{IX.4.15}$$

with the contours being circles of radii $r_i$ around the origin. It follows from the assumptions that $\big|\frac{\partial p}{\partial z_i}\big| > \mathrm{e}^{-a_i}$ for all $i$ and all $\boldsymbol{z} \in D$. Theorem IX.4.1 then follows from Lemma IX.4.3. □

*Proof of Theorem IX.4.2.* We use Eq. (IX.4.9) with $\Phi(\boldsymbol{z}) = \boldsymbol{z}^{\boldsymbol{k}}/\boldsymbol{\rho}(\boldsymbol{z})^{\boldsymbol{k}}$. We then repeat the proof of Theorem IX.4.1 without the term $p(\boldsymbol{z})$. This yields the bound

$$\begin{aligned} \Big| [\boldsymbol{\rho}^{\boldsymbol{n}}] \frac{\boldsymbol{z}(\boldsymbol{\rho})^{\boldsymbol{k}}}{\boldsymbol{\rho}^{\boldsymbol{k}}} \Big| &= \Big| [\boldsymbol{z}^{\boldsymbol{n}}] \frac{\boldsymbol{z}^{\boldsymbol{k}}}{\boldsymbol{\rho}(\boldsymbol{z})^{\boldsymbol{k}}} \frac{1}{\big(\frac{\partial p}{\partial \boldsymbol{z}}\big)^{\boldsymbol{n}}} \det M(\boldsymbol{z}) \Big| \\ &= \Big| [\boldsymbol{z}^{\boldsymbol{n}}] \frac{1}{\big(\frac{\partial p}{\partial \boldsymbol{z}}\big)^{\boldsymbol{n}+\boldsymbol{k}}} \det M(\boldsymbol{z}) \Big| \\ &< C \prod_{i\ge 1} \frac{\mathrm{e}^{a_i(n_i+k_i)}}{r_i^{n_i}}. \end{aligned} \tag{IX.4.16}$$

□

## IX.5 Connected and two-connected graphs

We now make a further assumption on the series $p(\boldsymbol{z})$ introduced in (IX.4.1): it is given by the (weighted) exponential generating function of coloured graphs. This choice is motivated by applications to statistical mechanics, which are discussed in

Section IX.6. If the weights satisfy a certain *block factorisation* property, the virial coefficients $c(\boldsymbol{n})$ can be expressed using two-connected graphs.

A coloured graph is a pair $(g, \boldsymbol{k})$ where $g$ is a graph and and $\boldsymbol{k}$ is a function $V(g) \to \mathbb{N}$ that assigns the colour $k_i \in \mathbb{N}$ to each $i \in V(g)$. Coloured connected graphs are pairs $(g, \boldsymbol{k})$ with $g$ a connected graph, coloured two-connected graphs are pairs $(g, \boldsymbol{k})$ with $g$ a two-connected graph. The results of this section can be formulated in the general framework of *labelled coloured combinatorial species* ( [MeNa93], [EhMe94]). Here they are presented in a self-contained way.

Let $w(g, \boldsymbol{k})$ be a weight function on coloured graphs. We assume that it is invariant under relabellings that preserve the colour: let $\sigma : V \to V$ be a bijection with the property that $k_{\sigma(i)} = k_i$ for all $i \in V$, and $g_\sigma$ the graph with vertices $V(g_\sigma) = V(g)$ and edges $E(g_\sigma) = \big\{\{\sigma(i), \sigma(j)\} : \{i, j\} \in E(g)\big\}$, then $w(g_\sigma, \boldsymbol{k}) = w(g, \boldsymbol{k})$. The weighted exponential generating function for connected graphs is defined by

$$\begin{aligned} C_w(\boldsymbol{z}) &= \sum_{n \geq 1} \frac{1}{n!} \sum_{\boldsymbol{k}=(k_1,\dots,k_n)} z_{k_1} \dots z_{k_n} \sum_{g \in \mathcal{C}[n]} w(g, \boldsymbol{k}) \\ &= \sum_{\boldsymbol{n}} \frac{\boldsymbol{z}^{\boldsymbol{n}}}{\boldsymbol{n}!} \sum_{\boldsymbol{g} \in \mathcal{C}[\boldsymbol{n}]} w(\boldsymbol{g}). \end{aligned} \tag{IX.5.1}$$

In the second line the sum is over multi-indices $\boldsymbol{n}$ with finitely many entries, and $\mathcal{C}[\boldsymbol{n}]$ denotes the set of coloured graphs $\boldsymbol{g} = (g, \boldsymbol{k}^{[\boldsymbol{n}]})$ with vertices $1, 2, \dots, |\boldsymbol{n}|$ and $\boldsymbol{k}^{[\boldsymbol{n}]}$ the colouring such that the first $n_1$ vertices have colour 1, the vertices $n_1+1, \dots, n_1+n_2$ have colour 2, etc. We shall refer to $\boldsymbol{k}^{[\boldsymbol{n}]}$ as the *canonical colouring*. The second expression for $C(\boldsymbol{z})$ is more elegant but the first expression turns out to be more practical. These formulæ should still be understood as formal series.

We suppose that $w(g, \boldsymbol{k})$ factorises with respect to the block decomposition of $g$. Recall that an *articulation point* of $g$ is a vertex $i \in \mathcal{V}$ such that the subgraph $g \setminus \{i\}$ is disconnected. (A two-connected graph is a connected graph without articulation point.)

**Theorem IX.5.1.** *Assume that $p(\boldsymbol{z}) = C_w(\boldsymbol{z})$ as above, that the weight function $w(g, \boldsymbol{k})$ satisfies $w(g, \boldsymbol{k}) = 1$ when $g$ has size $n = 1$, and*

$$w(g, \boldsymbol{k}) = \prod_{i=1}^{m} w(g_i, \boldsymbol{k}\big|_{V(g_i)})$$

*when $g$ has size $n \geq 2$, where $\{g_1, \dots, g_m\}$ is the block decomposition of $g$ and $\boldsymbol{k}\big|_{V(g_i)}$*

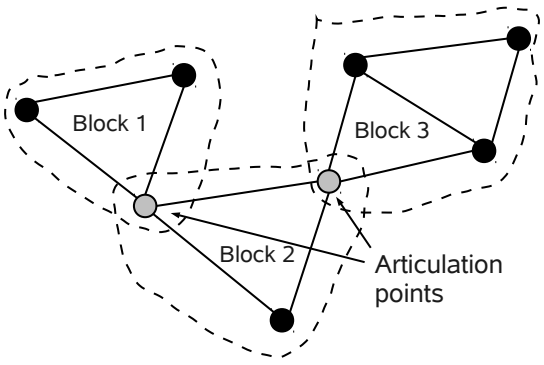



Figure IX.1: The block decomposition of a connected graph.

*is the restriction of the colouring* $\boldsymbol{k} : V(g) \to \mathbb{N}$ *to* $V(g_i)$*. Then*

$$p(\boldsymbol{z}) = \sum_{k\geq 1} \rho_k(\boldsymbol{z}) - \sum_{\boldsymbol{n}:\, |\boldsymbol{n}|\geq 2} \big(|\boldsymbol{n}| - 1\big) \frac{\boldsymbol{\rho}(\boldsymbol{z})^{\boldsymbol{n}}}{\boldsymbol{n}!} \sum_{\boldsymbol{g}\in\mathcal{B}[\boldsymbol{n}]} w(\boldsymbol{g})$$

*where* $\mathcal{B}[\boldsymbol{n}]$ *consists of the two-connected coloured graphs* $\boldsymbol{g} = (g, \boldsymbol{k}^{[\boldsymbol{n}]})$ *with canonical colouring* $\boldsymbol{k}^{[\boldsymbol{n}]}$ *and vertex set* $V(g) = \{1, 2, \ldots, |\boldsymbol{n}|\}$*.*

The proof is given at the end of this section. It uses the dissymmetry theorem of combinatorial structures, following [BLL98].

If $k \in \mathbb{N}$ denotes a colour, a $k$-rooted graph is a triplet $(g, \boldsymbol{k}, i)$ where $g$ is a graph with finite vertex set $V(g)$, $\boldsymbol{k} \in \mathbb{N}^{V(g)}$, $i \in V(g)$, with the property that the root $i$ has colour $k_i = k$. The weighted exponential generating function of $k$-rooted connected graphs is

$$\begin{aligned} C_w^{\bullet k}(\boldsymbol{z}) &= \sum_{n\geq 1} \frac{1}{n!} \sum_{\boldsymbol{k}=(k_1,\ldots,k_n)} z_{k_1} \ldots z_{k_n} \big|\{i : k_i = k\}\big| \sum_{g\in\mathcal{C}[n]} w(g, \boldsymbol{k}) \\ &= \sum_{\boldsymbol{n}} n_k \frac{\boldsymbol{z}^{\boldsymbol{n}}}{\boldsymbol{n}!} \sum_{\boldsymbol{g}\in\mathcal{C}[\boldsymbol{n}]} w(\boldsymbol{g}) \qquad \text{(IX.5.2)} \\ &= z_k \frac{\partial C}{\partial z_k}(\boldsymbol{z}). \end{aligned}$$

Let $\mathcal{C}$, $\mathcal{C}^{\bullet k}$ be the sets of connected (resp. $k$-rooted connected) coloured graphs with vertex set of the form $[n]$, $n \in \mathbb{N}$, and set and $\mathcal{C}^\bullet := \cup_{k\in\mathbb{N}} \mathcal{C}^{\bullet k}$. The associated exponential generating function is $C^\bullet(\boldsymbol{z}) = \sum_{k\in\mathbb{N}} C^{\bullet k}(\boldsymbol{z})$. We define $\mathcal{B}$, $\mathcal{B}^\bullet$ and their exponential generating functions $B(\boldsymbol{z})$, $B^\bullet(\boldsymbol{z})$ in a similar way, replacing "connected" with "two-connected".

Next, we describe the composition of connected and two-connected graphs. The set $\mathcal{B}(\{\mathcal{C}^{\bullet k}\}_{k\in\mathbb{N}})$ consists of coloured two-connected graphs whose vertices contain a rooted connected graph with the appropriate colour for the root. More precisely, an element $\boldsymbol{g} \in \mathcal{B}(\{\mathcal{C}^{\bullet k}\}_{k\in\mathbb{N}})$ of size $n$ is a triple $\boldsymbol{\gamma} = (\boldsymbol{k}, \gamma, (\gamma_i)_{i\in\mathcal{V}(\gamma)})$ consisting of

- A colour assignment $\boldsymbol{k} \in \mathbb{N}^n$.
- A two-connected graph $\gamma$ with vertex set $V(\gamma) \subset [n]$, $|V(\gamma)| \geq 2$.
- A family $(\gamma_i)_{i\in V(\gamma)}$ of connected graphs $\gamma_i$ such that $i \in V(\gamma_i)$, and the vertex sets form a partition $[n] = \cup_{i\in V(\gamma)} V(\gamma_i)$. Note that $(\gamma_i, \boldsymbol{k}|_{V(\gamma_i)}, i)$ is a $k_i$-rooted coloured connected graph.

With each $\boldsymbol{\gamma}$ we associate the connected graph $g = g(\boldsymbol{\gamma})$ with vertices $1, \dots, n$ and edge set $E(\gamma) \cup \big(\cup_{i\in V(\gamma)} E(\gamma_i)\big)$. We assign to the composite structure $\boldsymbol{\gamma}$ the weight $w(g(\boldsymbol{\gamma}), \boldsymbol{k})$ of the underlying connected coloured graph. We also introduce $\mathcal{B}^\bullet(\{\mathcal{C}^{\bullet k}\}_{k\in\mathbb{N}})$, which is as above but with the additional choice of a root in $V(\gamma)$.

**Lemma IX.5.2.** *Under the assumptions of Theorem IX.5.1, the weighted exponential generating functions of $\mathcal{B}(\{\mathcal{C}^{\bullet k}\}_{k\in\mathbb{N}})$ and $\mathcal{B}^\bullet(\{\mathcal{C}^{\bullet k}\}_{k\in\mathbb{N}})$ satisfy*

$$B(\{C^{\bullet k}(\boldsymbol{z})\}_{k\in\mathbb{N}}) = \sum_{\boldsymbol{\gamma}\in\mathcal{B}(\{\mathcal{C}^{\bullet k}\}_{k\in\mathbb{N}})} \frac{1}{n!} z_{k_1} \dots z_{k_n} w(g, \boldsymbol{k}),$$
$$B^\bullet(\{C^{\bullet k}(\boldsymbol{z})\}_{k\in\mathbb{N}}) = \sum_{\boldsymbol{\gamma}\in\mathcal{B}^\bullet(\{\mathcal{C}^{\bullet k}\}_{k\in\mathbb{N}})} \frac{1}{n!} z_{k_1} \dots z_{k_n} w(g, \boldsymbol{k}).$$

*Here, $n = n(\boldsymbol{g})$ is the size of $\boldsymbol{g}$, $\boldsymbol{k} = \boldsymbol{k}(\boldsymbol{g})$ is the colour assignment of the vertices $[n]$ and $g = g(\boldsymbol{\gamma})$ is the induced connected graph on the set of vertices $[n]$.*

The proofs are tedious but immediate: one sums over all components of $\boldsymbol{g}$, and uses the multinomial theorem so that the elements of the partition become independent. This is possible because of the factorisation property of the weights,

$$w(g(\boldsymbol{\gamma}), \boldsymbol{k}) = w(\gamma, \boldsymbol{k}|_{V(\gamma)}) \times \prod_{i\in V(\gamma)} w(\gamma_i, \boldsymbol{k}|_{V(\gamma_i)}). \tag{IX.5.3}$$

For a proof in the context of labelled coloured combinatorial species (but without weights), see [MeNa93] and [EhMe94].

Next, we state the dissymmetry theorem for coloured graphs.

**Theorem IX.5.3.** *We have*

$$\mathcal{C} + \mathcal{B}^\bullet(\{\mathcal{C}^{\bullet k}\}_{k\in\mathbb{N}}) = \mathcal{C}^\bullet + \mathcal{B}(\{\mathcal{C}^{\bullet k}\}_{k\in\mathbb{N}})$$

*in the sense that there is a size and weight preserving bijection between* $\mathcal{C}+\mathcal{B}^\bullet(\{\mathcal{C}^{\bullet k}\}_{k\in\mathbb{N}})$ *and* $\mathcal{C}^\bullet + \mathcal{B}(\{\mathcal{C}^{\bullet k}\}_{k\in\mathbb{N}})$.

*Proof.* There are two mappings $\phi : \mathcal{C}+\mathcal{B}^\bullet(\{\mathcal{C}^{\bullet k}\}_{k\in\mathbb{N}}) \to \mathcal{C}$ and $\psi : \mathcal{C}^\bullet+\mathcal{B}(\{\mathcal{C}^{\bullet k}\}_{k\in\mathbb{N}}) \to \mathcal{C}$, which associate to each graph in each of the sets above a unique connected graph. The idea is to informally 'forget' the extra structure afforded to us by the composite structures.

These two mappings are conveniently described in terms of their preimages (the structures corresponding to the same connected graph):

The preimage of $g$ under $\phi$ consists of the union of:

- The set containing the graph itself, $g \in \mathcal{C}$.
- The set of composite structures $(\boldsymbol{k}, \gamma, (\gamma_i)_{i\in V(\gamma)}, r) \in \mathcal{B}^\bullet(\{\mathcal{C}^{\bullet k}\}_{k\in\mathbb{N}})$ where $\gamma$ is one of the blocks $g_1, \dots, g_m$ of $g$, $r \in V(\gamma)$, and $(\gamma_i)$ are uniquely determined by $g$ and the choice of $\gamma$.

The preimage of $g$ under $\psi$ consists of the union of:

- The set of ordered pairs $(g, i)_{i\in[n]} \in \mathcal{C}^\bullet$, where the second entry indicates a root.
- The set of composite structures $(\boldsymbol{k}, \gamma, (\gamma_i)_{i\in V(\gamma)})$, where $\gamma$ is one of the blocks $g_1, \dots, g_m$ of $g$ and $(\gamma_i)$ are uniquely determined by $g$ and the choice of $\gamma$, as described above.

It is sufficient to prove that for every $\boldsymbol{g} = (g, \boldsymbol{k}) \in \mathcal{C}$, the preimages $\phi^{-1}(\{\boldsymbol{g}\})$ and $\psi^{-1}(\{\boldsymbol{g}\})$ have the same cardinality.

Let $\boldsymbol{g} = (g, \boldsymbol{k}) \in \mathcal{C}$. If $g$ has size one, then $\phi^{-1}(\{\boldsymbol{g}\}) = \{\boldsymbol{g}\} \subset \mathcal{C}$ and $\psi^{-1}(\{\boldsymbol{g}\}) = \{(g, \boldsymbol{k}, 1)\} \subset \mathcal{C}^\bullet$, and both preimages have cardinality 1. If $g$ has size $n \geq 2$, let $\{g_1, \dots, g_m\}$ be the block decomposition of $g$.

The preimage under $\phi$ has cardinality

$$1 + \sum_{i=1}^{m} |V(g_i)|. \tag{IX.5.4}$$

The sum gives the possible roots for each block considered in turn.

The preimage under $\psi$ has cardinality

$$n + m. \tag{IX.5.5}$$

The first "$n$" corresponds to the number of ways to choose the root when the preimage is in $\mathcal{C}^\bullet$, and $m$ is the number of composite structures $\boldsymbol{\gamma} \in \mathcal{B}^\bullet(\{\mathcal{C}^{\bullet k}\}_{k\in\mathbb{N}})$ with $g(\boldsymbol{\gamma}) = g$.

There remains to show that the block decomposition of every $g \in \mathcal{C}[n]$ satisfies

$$1 + \sum_{i=1}^{m} |V(g_i)| = n + m, \tag{IX.5.6}$$

or equivalently

$$\sum_{i=1}^{m} \big(|V(g_i)| - 1\big) = n - 1. \tag{IX.5.7}$$

This can be seen by induction. This clearly holds when $m = 1$ and $\mathcal{V}(g_1) = [n]$ (this corresponds to $g$ being two-connected). Now suppose that $g \in \mathcal{C}[n]$ has $m$ blocks of size $n_1, \ldots, n_m$. Consider the bipartite graph $t$ whose vertex set consists of the blocks $g_1, \ldots, g_m$ and articulation points $a_1, \ldots, a_\ell$ of $g$, and edges $\{\{g_i, a_j\} : a_j \in g_i\}$. The graph $t$ is known [BLL98] to be a tree and called the *block cut-point tree* of $g$. Let $v \in \mathcal{V}(t)$ be a leaf of $t$. Then $v$ is a vertex belonging to exactly one edge and must be a block $v = g_i$ containing exactly one articulation point $a$ of $g$.

Thus there is a block containing precisely one articulation point of $g$, and without loss of generality we take this to be the $m$th block. We remove from $g$ all edges of the block $g_m$ and all vertices of $g_m$, except the articulation point $a$. We now have a graph with $m - 1$ blocks and so we have $\sum_{i=1}^{m-1}(n_i - 1) = n - n_m$ by induction. Therefore,

$$\sum_{i=1}^{m}(n_i - 1) = n - n_m + n_m - 1 = n - 1. \tag{IX.5.8}$$

□

*Proof of Theorem IX.5.1.* Lemma IX.5.2 and the dissymmetry theorem imply that the exponential generating functions satisfy

$$C(\boldsymbol{z}) + B^\bullet(\{C^{\bullet k}(\boldsymbol{z})\}_{k\in\mathbb{N}}) = C^\bullet(\boldsymbol{z}) + B(\{C^{\bullet k}(\boldsymbol{z})\}_{k\in\mathbb{N}}). \tag{IX.5.9}$$

To directly compare with Theorem IX.5.1, we write this as:

$$C(\boldsymbol{z}) = C^{\bullet}(\boldsymbol{z}) + B(\{C^{\bullet k}(\boldsymbol{z})\}_{k\in\mathbb{N}}) - B^{\bullet}(\{C^{\bullet k}(\boldsymbol{z})\}_{k\in\mathbb{N}}). \tag{IX.5.10}$$

We have $p(\boldsymbol{z}) = C(\boldsymbol{z})$ and $\rho_k(\boldsymbol{z}) = C^{\bullet k}(\boldsymbol{z})$, and the theorem follows. □

We conclude this section with two remarks. The first remark is that under the assumptions of Theorem IX.5.1, we also have a formula for the expansion of the chemical potential $\log z_k$ in terms of the density,

$$\log z_k = \log \rho_k(\boldsymbol{z}) - \frac{\partial B}{\partial \rho_k}\big(\boldsymbol{\rho}(\boldsymbol{z})\big), \quad B(\boldsymbol{\rho}) = \sum_{\boldsymbol{n}:\,|\boldsymbol{n}|\geq 2} \frac{\boldsymbol{\rho}^{\boldsymbol{n}}}{\boldsymbol{n}!} \sum_{\boldsymbol{g}\in\mathcal{B}[\boldsymbol{n}]} w(\boldsymbol{g}). \tag{IX.5.11}$$

This follows from the relation

$$\rho_k(\boldsymbol{z}) = C^{\bullet k}(\boldsymbol{z}) = z_k \exp\Big(\frac{\partial B}{\partial \rho_k}\big(\boldsymbol{\rho}(\boldsymbol{z})\big)\Big) \tag{IX.5.12}$$

[Far10]; $\partial_k B$ is the generating function for two-connected graphs whose root is a "ghost" of colour $k$.

The second remark is that Theorem IX.5.1 is not limited to connected and two-connected graphs, but holds for pairs of combinatorial structures with a similar composition structure – as Leroux puts it, for "various tree-like structures" [Ler04]. A well-known example is the dissymmetry theorem for trees [BLL98], which can be adapted to coloured trees with colour-dependent weights and constraints. This is interesting because Good's original motivation for his multi-variable version of the Lagrange inversion came from branching processes in probability and combinatorics of trees [Good60, Good65]. Further notes on the dissymmetry theorem are found in Section VI.3.

## IX.6 Classical gas of rigid molecules

The context of the application to a classical gas of rigid molecules is along the lines of the paper [PoUe09]. There is a key issue here in not being able to achieve the necessary uniform lower bounds for the logarithm of the possibly complex-valued functions, which could give continuous internal degrees of freedom. The term rigid indicates that the pair potential between two particles is independent of further degrees of freedom than particle type, such as particle orientation or spin.

We now describe a physical system that fits the theory of Sections IX.4 and IX.5. It consists of a gas of molecules that are assumed to be rigid. Let $\Lambda \subset \mathbb{R}^d$ be

the domain, which we take as a cube in $\mathbb{R}^d$ with periodic boundary conditions. We let $V$ denote its volume. A molecule is represented by

$$X = (k, x, \phi), \tag{IX.6.1}$$

where $k \in \mathbb{N}$ denotes the species, $x \in \Lambda$ denotes the position, and $\phi \in \Phi := S^{d-1} = \{\phi \in \mathbb{R}^d \mid \|\phi\| = 1\}$ denotes the orientation. Interactions are given by a function $U(X_1, X_2)$ that takes values in $\mathbb{R} \cup \{+\infty\}$. Let

$$\zeta(X_1, X_2) = \mathrm{e}^{-U(X_1, X_2)} - 1. \tag{IX.6.2}$$

We take periodic boundary conditions, i.e., we assume that if $\Lambda = [0, L]^d$, and $y - y' \in L\mathbb{Z}^d$, then $U((x, k, \phi), (y, \ell, \psi)) = U((x, k, \phi), (y', \ell, \psi))$. We make three assumptions on the interactions. The first one is about symmetries, the second one is the stability condition that ensures the existence of the thermodynamic limit, and the last one implies that we consider a regime of low density or high temperatures.

**Assumption 10.** *The potential function $U$ satisfies*

- *Symmetry: $U(X_1, X_2) = U(X_2, X_1)$.*
- *Translation invariance: If $X + a$ denotes the molecule translated by $a \in \Lambda$, i.e., with position $x + a$, then $U(X_1 + a, X_2 + a) = U(X_1, X_2)$.*
- *Rotation invariance: If $RX$ denotes the molecule rotated by the orthogonal matrix $R$, i.e., with orientation $R\phi$, then $U(RX_1, RX_2) = U(X_1, X_2)$.*

The *partition function* of the system is

$$Z_\Lambda(\boldsymbol{z}) = \sum_{n \geq 0} \frac{1}{n!} \sum_{\boldsymbol{k} \in \mathbb{N}^n} z_{k_1} \dots z_{k_n} \int_{\Lambda^n} \mathrm{d}x_1 \dots \mathrm{d}x_n \int_{\Phi^n} \mathrm{d}\phi_1 \dots \mathrm{d}\phi_n \exp\Big\{ - \sum_{1 \leq i < j \leq n} U(X_i, X_j) \Big\}. \tag{IX.6.3}$$

The term $n = 0$ is understood to be equal to 1, and $\mathrm{d}\phi$ is the unique rotationally invariant measure on $\Phi$ with $\int_\Phi \mathrm{d}\phi = 1$.

Given a graph $g \in \mathcal{C}[n]$, we define the weight function $w_\Lambda(g, \boldsymbol{k})$ to be

$$w_\Lambda(g, \boldsymbol{k}) = \frac{1}{V} \int_{\Lambda^n} \mathrm{d}x_1 \dots \mathrm{d}x_n \int_{\Phi^n} \mathrm{d}\phi_1 \dots \mathrm{d}\phi_n \prod_{\{i,j\} \in \mathcal{E}(g)} \zeta(X_i, X_j), \tag{IX.6.4}$$

where

$$\zeta(X_i, X_j) = \mathrm{e}^{-U(X_i, X_j)} - 1. \tag{IX.6.5}$$

The empty product is set to be equal to 1, so that graphs of size 1 have weight $V^{-1}\int_\Lambda \mathrm{d}x \int_\Phi \mathrm{d}\phi = 1$. By a standard cluster expansion, or by the exponential formula of combinatorial structures, we have

$$Z_\Lambda(\boldsymbol{z}) = \exp\Big\{V \sum_{n\geq 1} \frac{1}{n!} \sum_{\boldsymbol{k}\in\mathbb{N}^n} z_{k_1}\dots z_{k_n} \sum_{g\in\mathcal{C}_n} w_\Lambda(g,\boldsymbol{k})\Big\}. \tag{IX.6.6}$$

The partition function is related to the finite-volume pressure by $Z_\Lambda = \mathrm{e}^{Vp_\Lambda}$. We then define

$$p_\Lambda(\boldsymbol{z}) = C(\boldsymbol{z}), \tag{IX.6.7}$$

where $C(\boldsymbol{z})$ is the exponential generating function of connected graphs given in (IX.5.1) with the weights $w_\Lambda(g,\boldsymbol{k})$ in (IX.6.4). The goal is to show that the assumptions of Theorem IX.4.1 hold true uniformly in the volume $V$.

**Lemma IX.6.1.** *Under Assumption 10, the weight function of (IX.6.4) satisfies the block factorisation*

$$w_\Lambda(g,\boldsymbol{k}) = \prod_{i=1}^m w_\Lambda(g_i,\boldsymbol{k}).$$

It is not too hard to check that factorisation holds when the graph is cut at any articulation point, and the lemma follows. It should be stressed that Lemma IX.6.1 fails when the molecules are not assumed to be rigid. We emphasise here that the dissymmetry theorem always holds, but it is a case of the weights working out correctly. For graphs which block-factorise we have the nice interpretation given in this section, otherwise some ideas on the weight modification required are presented in section IX.8. Next, the stability condition.

**Assumption 11.** *There exists a nonnegative constant $b$ such that for all $n$ and all $X_1,\dots,X_n$, we have*

$$\prod_{1\leq i<j\leq n} \big|1+\zeta(X_i,X_j)\big| \leq \prod_{i=1}^n \mathrm{e}^{bk_i}. \tag{IX.6.8}$$

*In addition, we also assume that for all $X$ and $Y$ of species $k$ and $\ell$, we have*

$$\big|1+\zeta(X,Y)\big| \leq \mathrm{e}^{b\min(k,\ell)}. \tag{IX.6.9}$$

The next and last assumption is the "Kotecký-Preiss criterion" that guarantees that the interactions and the weights are small.

**Assumption 12.** *There exist positive numbers $R_1, R_2, \dots$ and a constant $a$ such that for all $X = (k, x, \phi)$,*

$$\sum_{k' \in \mathbb{N}} R_{k'} \, \mathrm{e}^{(a+3b)k'} \int_{\mathbb{R}^d} \mathrm{d}x' \int_{\Phi} \mathrm{d}\phi' |\zeta(X, X')| \leq ak, \tag{IX.6.10}$$

*where $b$ is given in Assumption 11. In addition, we also assume that*

$$\sum_{k' \in \mathbb{N}} R_{k'} \, \mathrm{e}^{(a+3b)k'} < \infty. \tag{IX.6.11}$$

**Theorem IX.6.2.** *Let $p_\Lambda(\boldsymbol{z}) = C(\boldsymbol{z})$ and suppose that Assumptions 10–12 hold true. Then*

*(a) $p_\Lambda(\boldsymbol{z})$ converges absolutely in the polydisc $D = \{\boldsymbol{z} \in \mathbb{C}^{\mathbb{N}} : |z_i| < R_i \; \forall i \in \mathbb{N}\}$.*

*(b) $\left| \log \dfrac{\partial p_\Lambda}{\partial z_k}(\boldsymbol{z}) \right| < ak$ for all $\boldsymbol{z} \in D$ and all $k \in \mathbb{N}$.*

The main consequence of this theorem is that Theorem IX.4.1 applies, hence the existence of a domain of densities with absolute convergence of the virial expansion.

*Proof.* The setting of [PoUe09] applies directly here. The measure space of "polymers" $(\mathbb{X}, \mu)$ in [PoUe09] is presently given by $\mathbb{X} = \mathbb{N} \times \Lambda \times \Phi$ with $\mu$ the measure such that

$$\int_{\mathbb{X}} f(X) \mathrm{d}\mu(X) = \sum_{k \in \mathbb{N}} z_k \int_{\Lambda} \mathrm{d}x \int_{\Phi} \mathrm{d}\phi f(k, x, \phi) \tag{IX.6.12}$$

for arbitrary integrable function $f$ on $\mathbb{X}$.

The conditions of [PoUe09] are fulfilled — our Assumption 12 being slightly stronger with $3b$ instead of $2b$, but it will be needed in the proof of (b). From Theorem 2.1 in [PoUe09] we have that for every $X_1 = (k_1, x_1, \phi_1)$, and every $\boldsymbol{z} \in D$,

$$\begin{aligned} \sum_{n \geq 2} \frac{1}{(n-1)!} \sum_{k_2, \dots, k_n \in \mathbb{N}} |z_{k_2}| \dots |z_{k_n}| \int_{\mathbb{R}^d} \mathrm{d}x_2 \dots \int_{\mathbb{R}^d} \mathrm{d}x_n \\ \int_{\Phi} \mathrm{d}\phi_2 \dots \int_{\Phi} \mathrm{d}\phi_n \Big| \sum_{g \in \mathcal{C}[n]} \prod_{\{i,j\} \in E(g)} \zeta(X_i, X_j) \Big| \leq (\, \mathrm{e}^{ak_1} - 1) \, \mathrm{e}^{2bk_1} \, . \end{aligned} \tag{IX.6.13}$$

In particular, the Taylor series of the pressure $p_\Lambda(\boldsymbol{z})$ is absolutely convergent in $D$, uniformly in $\Lambda$.

For (b) we need to control the logarithm of the derivative of $p_\Lambda$. It is not entirely straightforward as we need both lower and upper bounds for $\frac{\partial}{\partial z_k} p_\Lambda$. We

have

$$\frac{\partial p_\Lambda}{\partial z_k} = \frac{1}{V}\frac{\partial}{\partial z_k}\log Z_\Lambda(\boldsymbol{z}) = \frac{1}{V}\frac{1}{Z_\Lambda(\boldsymbol{z})}\frac{\partial Z_\Lambda}{\partial z_k}(\boldsymbol{z}). \tag{IX.6.14}$$

From the definition (IX.6.3) of the partition function, we get

$$\frac{\partial Z_\Lambda}{\partial z_k}(\boldsymbol{z}) = V\sum_{n\geq 1}\frac{1}{(n-1)!}\sum_{k_2,\dots,k_n\geq 1} z_{k_1}\dots z_{k_n}\int_{\Lambda^{n-1}}\mathrm{d}x_2\dots\mathrm{d}x_n\int_{\Phi^{n-1}}\mathrm{d}\phi_2\dots\mathrm{d}\phi_n \\ \exp\Big\{-\sum_{j=2}^{n}U(X,X_j) - \sum_{2\leq i<j\leq n}U(X_i,X_j)\Big\}. \tag{IX.6.15}$$

We set $X=(k,0,0)$. The formula holds because of translation and rotation invariance, and because $\int \mathrm{d}\phi_1 = 1$. We observe that $\frac{\partial Z_\Lambda}{\partial z_k}$ is a partition function where each molecule $X_j$ gets the extra factor $\mathrm{e}^{-U(X,X_j)}$. We can again perform a cluster expansion or use the exponential formula of combinatorial structures. It is indeed convergent thanks to (IX.6.9). We get

$$\frac{\partial Z_\Lambda}{\partial z_k}(\boldsymbol{z}) = V\exp\Big\{\sum_{n\geq 1}\frac{1}{n!}\sum_{k_1,\dots,k_n\in\mathbb{N}} z_{k_1}\dots z_{k_n}\int_{\Lambda^n}\mathrm{d}x_1\dots\mathrm{d}x_n\int_{\Phi^n}\mathrm{d}\phi_1\dots\mathrm{d}\phi_n \\ \prod_{j=1}^{n}\mathrm{e}^{-U(X,X_j)}\sum_{g\in\mathcal{C}[n]}\prod_{\{i,j\}\in E(g)}\zeta(X_i,X_j)\Big\}. \tag{IX.6.16}$$

This allows us to combine it with the cluster expansion of $Z_\Lambda$ in (IX.6.14) and we get

$$\frac{\partial p_\Lambda}{\partial z_k}(\boldsymbol{z}) = \exp\Big\{\sum_{n\geq 1}\frac{1}{n!}\sum_{k_1,\dots,k_n\in\mathbb{N}} z_{k_1}\dots z_{k_n}\int_{\Lambda^n}\mathrm{d}x_1\dots\mathrm{d}x_n\int_{\Phi^n}\mathrm{d}\phi_1\dots\mathrm{d}\phi_n \\ \Big(\prod_{j=1}^{n}\big(1+\zeta(X,X_j)\big)-1\Big)\sum_{g\in\mathcal{C}[n]}\prod_{\{i,j\}\in E(g)}\zeta(X_i,X_j)\Big\}. \tag{IX.6.17}$$

Next we use the identity

$$\prod_{j=1}^{n}\big(1+\zeta(X,X_j)\big)-1 = \Big[\prod_{j=1}^{n-1}\big(1+\zeta(X,X_j)\big)-1\Big]\big(1+\zeta(X,X_n)\big)+\zeta(X,X_n). \tag{IX.6.18}$$

It allows to prove by induction that

$$\Big|\prod_{j=1}^{n}\big(1+\zeta(X,X_j)\big)-1\Big| \leq \mathrm{e}^{b\sum_{j=1}^{n}k_j}\sum_{j=1}^{n}|\zeta(X,X_j)|. \tag{IX.6.19}$$

The integrand of (IX.6.17) is then less than

$$\begin{aligned}\sum_{k_1\in\mathbb{N}} |z_{k_1}|\,\mathrm{e}^{bk_1} \int_\Lambda \mathrm{d}x_1 \int_\Phi \mathrm{d}\phi_1 |\zeta(X,X_1)| \Big(1 + \sum_{n\geq 2} \frac{1}{(n-1)!} \sum_{k_2,\dots,k_n\in\mathbb{N}} |z_2\dots z_n|\,\mathrm{e}^{b\sum_{i=2}^n k_i} \\ \int_{\Lambda^{n-1}} \mathrm{d}x_2\dots\mathrm{d}x_n \int_{\Phi^{n-1}} \mathrm{d}\phi_2\dots\mathrm{d}\phi_n \Big| \sum_{g\in\mathcal{C}[n]} \prod_{\{i,j\}\in E(g)} \zeta(X_i,X_j)\Big|\Big) \\ \leq \sum_{k_1\in\mathbb{N}} |z_{k_1}|\,\mathrm{e}^{(a+3b)k_1} \int_\Lambda \mathrm{d}x_1 \int_\Phi \mathrm{d}\phi_1 |\zeta(X,X_1)| \\ \leq ak.\end{aligned} \tag{IX.6.20}$$

We bounded the parenthesis by $\mathrm{e}^{(a+2b)k_1}$ using (IX.6.13). The last inequality follows from Assumption 12. □

## IX.7 The point of view of the inverse function theorem

The formulation of Theorem IX.4.1 is geared towards the virial expansion in statistical mechanics. The theorem in itself is, however, purely analytic. In this section we rephrase it as a type of inverse function theorem and discuss its relation to traditional inverse function theorems.

Let $\big(F_k(\boldsymbol{w})\big)_{k\in\mathbb{N}}$ be a family of power series in the complex variables $w_j$ ($j\in\mathbb{N}$) such that $F_k(0)\neq 0$, for all $k\in\mathbb{N}$. The goal is to invert the system of equations $u_k = w_k F_k(\boldsymbol{w})$. On the level of formal power series, the inversion is always possible: there is a unique family of power series $G_k(\boldsymbol{u})$, $k\in\mathbb{N}$, such that the inverse is given by $w_k(\boldsymbol{u}) = u_k G_k(\boldsymbol{u})$, i.e., for all $k\in\mathbb{N}$, we have $w_k(\boldsymbol{u})F_k(\boldsymbol{w}(\boldsymbol{u})) = u_k$, as an identity of formal power series.

**Theorem IX.7.1.** *Assume that there exist $0 < r_i < R_i$ and $a_i \geq 0$, $i = 1, 2, \dots$, such that*

- *The series $F_k(\boldsymbol{w})$, $k\in\mathbb{N}$, converge absolutely in the polydisc $D = \{\boldsymbol{w}\in\mathbb{C}^{\mathbb{N}} \mid \forall i\in\mathbb{N}: |w_i| < R_i\}$.*
- *$|\log F_i(\boldsymbol{w})| < a_i$ for all $i\geq 1$ and all $\boldsymbol{w}\in D$.*
- *$\displaystyle\sum_{i\geq 1}\sqrt{\frac{r_i}{R_i}} < \infty$ and $\displaystyle\sum_{i\geq 1}\frac{r_i a_i^2}{R_i} < \infty$.*

*Then there exists a constant $C < \infty$ (which depends on the $r_i$, $R_i$, $a_i$, but not on*

*$\boldsymbol{n}$) such that for all $k \in \mathbb{N}$,*

$$\big|[\boldsymbol{u}^{\boldsymbol{n}}]G_k(\boldsymbol{u})\big| \le C\,\mathrm{e}^{a_k} \prod_{i\ge 1}\Big(\frac{\mathrm{e}^{a_i}}{r_i}\Big)^{n_i}.$$

The proof is similar to the proofs of Theorems IX.4.1 and IX.4.2 and it is omitted. In order to better understand the analytic structure of the theorem, it is convenient to introduce Banach spaces of complex sequences. To simplify matters, suppose that the first two assumptions of the theorem hold with $a_k = ak$, where $a > 0$ is some $k$-independent constant, and $R_k = \exp(-ak)$. Situations of this type occur in the context of cluster expansions. Choose $r_k := k^{-4}\exp(-ak)$. Define the weighted $\ell^1$-norms $\|\boldsymbol{w}\|_{p,q} := \sum_{k\in\mathbb{N}} k^p \exp(qak)|w_k|$. Let us define

$$B_{p,q}(\varepsilon) := \{\boldsymbol{w} \in \mathbb{C}^{\mathbb{N}} : \|\boldsymbol{w}\|_{p,q} < \varepsilon\}. \tag{IX.7.1}$$

Fix $\varepsilon \in (0,1)$ small enough so that $C\varepsilon/(1-\varepsilon) < 1$ and set $c_\varepsilon := (1-\varepsilon)^{-1}\exp(-\varepsilon)$. As a consequence of Theorem IX.7.1 we get the well-defined functions

$$\begin{aligned} \boldsymbol{f} &: B_{0,1}(1) \to B_{0,0}(1) & \boldsymbol{g} &: B_{4,2}(\varepsilon) \to B_{4,1}\big(C\varepsilon/(1-\varepsilon)\big) \subset B_{0,1}(1) \\ &\boldsymbol{w} \mapsto \big(w_k F_k(\boldsymbol{w})\big)_{k\in\mathbb{N}} & &\boldsymbol{u} \mapsto \big(u_k G_k(\boldsymbol{u})\big)_{k\in\mathbb{N}}. \end{aligned} \tag{IX.7.2}$$

Then for every $\boldsymbol{u} \in B_{4,2}(\varepsilon)$, we have

$$\boldsymbol{f}\big(\boldsymbol{g}(\boldsymbol{u})\big) = \boldsymbol{u}. \tag{IX.7.3}$$

It is instructive to compare this with traditional inverse function theorems: suppose that $\boldsymbol{f}$, considered as map from the Banach space with norm $\|\cdot\|_{4,1}$ to the space with norm $\|\cdot\|_{4,2}$, was Fréchet-differentiable with invertible derivative in a neighbourhood of the origin. Then we would get the existence of open neighbourhoods of the origin such that $\boldsymbol{f}$ is a bijection between these neighbourhoods [Zei95]. Our result yields a weaker conclusion: we have a bijection between $B_{4,2}(\varepsilon)$ and $\boldsymbol{g}(B_{4,2}(\varepsilon))$, but in general the latter set needs not be open with respect to $\|\cdot\|_{4,1}$. The reason is that Theorem IX.7.1 operates under conditions that are weaker than those of the inverse function theorem: for infinitely many variables, the existence of continuous partial derivatives $\partial\boldsymbol{f}/\partial w_k$ does not imply that $\boldsymbol{f}$ is Fréchet-differentiable — we do not know whether the Jacobi matrix $(\partial f_\ell/\partial w_k)$ represents a bounded operator. Our condition $\exp(-a_k) \le |f_k(\boldsymbol{w})/w_k| \le \exp(a_k)$ replaces the traditional condition that the derivative and its inverse are bounded operators between Banach spaces.

## IX.8 The Dissymmetry Theorem without Block Factorisation

Since the dissymmetry theorem works on the level of structures, relating connected graphs to two-connected graphs, it is possible to consider whether we may understand it applying to cases where block-factorisation is not possible. Indeed, the key problem that is faced when there is no possibility of block-factorisation is what weights should be used for the two-connected graphs in the composition of structures. In this chapter, we have seen the generalisation to an arbitrary but countable number of colours to decorate the species. This resolves to decorating the graphs with colours to represent different species. The question is whether we can generalise this to continuous labels and achieve functionals to indicate the weights of the various structures in this example.

The set up is to have the activity dependent on position $z(x_i)$. In this case, we have a generating functional for the pressure expansion:

$$\beta P = \sum_{n=1}^{\infty} \frac{1}{n!} \sum_{g \in \mathcal{C}[n]} \frac{1}{|\Lambda|} \prod_{i \in V(g)} (\in z(x_i)\, \mathrm{d}x_i) \prod_{\{j,k\} \in E(g)} f_{jk} \tag{IX.8.1}$$

In this case, we may obtain a position dependent density as a functional derivative of pressure:

$$\rho(x) := z(x) \frac{\delta}{\delta z(x)} \beta P \tag{IX.8.2}$$

The key idea is then to observe pressure as the generating functional of connected graphs and density as the rooted version. The Dissymmetry Theorem may then be integrated over the continuous label set for the position dependence in density, in analogy with the coloured species, where the label set is discrete. This leads to making the following conjecture.

**Conjecture IX.8.1** (Functional Version of the Dissymmetry Theorem)**.** *If density is understood as the spatial distribution $\rho(x)$ for $x \in \mathbb{R}^D$ as above and $V_k(1, \cdots, k)$ is the Husimi function, then we have the following expansion of pressure in terms of density:*

$$\beta P = \int_{\mathbb{R}^D} \rho(x)\, \mathrm{d}^D x - \sum_{n \geq 2} \frac{n-1}{n!} \int_{\mathbb{R}^D} \cdots \int_{\mathbb{R}^D} V_n(1, \cdots, n) \prod_{i=1}^{n} \rho(x_i)\, \mathrm{d}^D x_i \tag{IX.8.3}$$

The intention here is to indicate that this may be a possibility and it is

something to look into further with the details of models which break the block-multiplicativity proeprty that is required for the simpler explanation.

## IX.9 Conclusions & Open Questions

It is possible to use Lagrange-Good inversion to understand the analytic properties for multispecies inversion. The tools of Analysis are not so readily available in this context. This algebraic combinatorial identity is therefore a valuable alternative. Furthermore, the interpretation of the Lagrange-Good inversion used here is helpful in many further contexts such as in renewal processes as is indicated in [JaTs13].

There are key issues that still need to be addressed about how to extend the conditions required to more complex models rather than only rigid polymers. Indeed, it is necessary to try and understand how continuous internal degrees of freedom effects the ability to bound the pressure logarith from below.

The dissymmetry theorem may also be generalised in this context, giving block multiplicativity of the weight function as a sufficient condition to achieve it. It is still unclear what happens in a case where this is not possible and can provide a useful avenue of further research.

# Chapter X

# Conclusions & Open Questions

The thesis gives an improvement of the bounds available for the virial coefficients and consequently the radius of convergence of the virial expansion. The bounds are presented as:

$$|c_n| \leq \frac{1}{n}\left( C(\beta)e^{4\beta B}\frac{W\left(\frac{e}{1+e^{2\beta B}}\right)}{\left(W\left(\frac{e}{1+e^{2\beta B}}\right)-1\right)^2}\right)^{n-1} \tag{X.0.1}$$

where $c_n$ are the virial coefficients and $W$ is the Lambert $W$-function. They give as a consequence the lower bound on the radius of convergence of the virial expansion $\mathcal{R}_{\text{vir}}$:

$$\mathcal{R}_{\text{vir}} \geq C(\beta)^{-1}e^{-4\beta B}\frac{\left(W\left(\frac{e}{1+e^{2\beta B}}\right)-1\right)^2}{W\left(\frac{e}{1+e^{2\beta B}}\right)} \tag{X.0.2}$$

A further bound is also made possible through the alternative tree bounds given by Procacci [Pro07] and Poghosyan Ueltschi [PoUe09], coming from the work of Brydges and Federbush [BrFe78]. This involves the modified notion of temperedness $R(\beta)$ and gives coefficient bounds:

$$|c_n| \leq \frac{1}{n}R(\beta)^{n-1}\left(\frac{W\left(\frac{e^{\beta B+1}}{1+e^{\beta B}}\right)}{\left(W\left(\frac{e^{\beta B+1}}{1+e^{\beta B}}\right)-1\right)^2}\right)^{n-1} \tag{X.0.3}$$

and a lower bound on the radius of convergence:

$$\mathcal{R}_{\text{vir}} \geq R(\beta)^{-1}\frac{\left(W\left(\frac{e^{\beta B+1}}{1+e^{\beta B}}\right)-1\right)^2}{W\left(\frac{e^{\beta B+1}}{1+e^{\beta B}}\right)} \tag{X.0.4}$$

It is still left open to consider models for which the latter condition is advantageous over the former.

In addition to improved bounds and original bounds for the virial expansion an original proof is given for the two identities:

$$\sum_{g\in\mathcal{B}[n]} (-1)^{e(g)} = -(n-2)! \tag{X.0.5}$$

$$\sum_{g\in\mathcal{B}[n]} (-1)^{e(g)} \operatorname{Vol}(\Pi_g) = -n(n-2)! \tag{X.0.6}$$

where the polytope $\Pi_g$ is defined as:

$$\Pi_g := \{(\mathbf{x})_{[2,n]} \in \mathbb{R}^{n-1} |\ |\mathbf{x}_i - \mathbf{x}_j| < 1\ \forall \{i,j\} \in E(g)\ \mathbf{x}_1 = 0\} \tag{X.0.7}$$

This involves interpreting the coefficients on the right hand side of these equations as a combinatorial object fixed by an involution on the two-connected graphs. The interpretation involves and increasing tree on $[n-1]$ with the vertex $n$ adjacent to every other vertex in the former case and all cases of cyclic relabellings (that is a map $i \mapsto i + k(\mathrm{mod} n)$) of these graphs in the latter case, with each relabelling considered distinct.

The graph tree expansions have deep connections to this type of involution as presented in the paper by Bernardi [Ber08]. The extensions of Penrose partitionality to matroids in Sokal's work on the arithmetic Tutte polynomial, introducing the notion of internally and externally active edges, as an effective way to gather cancellations in the partition functions is hoped to be able to be extended to two-connected graphs. There are some inherent difficulties here such as the fact that two-connected graphs do not have minimal graphs of the same size and the minimal graphs appear to diverge. Furthermore, a partition cannot occur in the simple way in which it does for connected graphs. An open question is to still see if the combinatorial understanding of the two simple statistical models presented in Chapter III can have an impact in formulating a method to achieving such an expansion.

Finally, multispecies virial expansions are made possible through Lagrange Good inversion. Conditions are formulated for the convergence of the virial expansion, based upon the behaviour of the cluster coefficients in this case. The Dissymmetry Theorem allows one to interpret the virial coefficients as coloured two-connected graphs in the case of block-multiplicative weight functions. Furthermore, it is applicable to a model of rigid polymers. However, there are difficulties in extending this to particles with continuous internal degrees of freedom.

Open questions from the multispecies consideration are whether there are further models to apply the analysis to and whether there is any possibility of understanding the requirements of the two-connected graph interpretation. There are also questions on the placement of this idea within Analysis, in that there is an 'inverse function theorem' for functions which need not be Fréchet differentiable. Furthermore, the concept of Lagrange Good inversion provides a wide scope of applications and interpretations within the literature. It would be useful to understand the connections between these.

# Bibliography


[Abd03a] A. Abdesselam, *A physicist's proof of the Lagrange-Good multivariable inversion formula*, Journal of Physics A 36, 9471–9477 (2003)

[Abd03b] A. Abdesselam, *The Jacobian conjecture as a problem of perturbative quantum field theory* In Annales Henri Poincaré, 4, 2, 199–215. Birkhäuser Verlag, (2003)

[AbRi94] A. Abdesselam, V. Rivasseau, *Trees, forests and jungles: a botanical garden for cluster expansions* in Constructive Physics, Proceedings (1994)

[BaLe64] S. Baert and J. L. Lebowitz, *Convergence of fugacity expansion and bounds on molecular distributions for mixtures*, Journal of Chemical Physics 40, 3474–3478 (1964)

[BLP66] J. A. Barker, P. J. Leonard, and A. Pompe *Fifth virial coefficients* The Journal of Chemical Physics, 44, 4206 (1966)

[Bat84] G. Battle, *A new combinatoric estimate for cluster expansions* Communications in Mathematical Physics, 94, 133–139 (1984)

[BaFe82] G. Battle, and P. Federbush, *A phase cell cluster expansion for Euclidean field theories* Annals of Physics, 142, 95–139 (1982)

[BaFe83] G. Battle, and P. Federbush, *A phase cell cluster expansion for a hierarchical $\phi^4_3$ model* Commununications in Mathematical Physics, 88, 263–293 (1983)

[BaFe84] G. Battle and P. Federbush *A note on cluster expansions, tree-graph identities, extra 1/N! factors!!!* Letters in Mathematical Physics, 8, 55–57 (1984)

[BeGo71] E. A. Bender, J. R. Goldman, *Enumerative uses of generating functions* Indiana University Mathematics Journal 20, 753, (1971)

[BLL98] F. Bergeron, G. Labelle, and P. Leroux, *Combinatorial Species and Tree-like Structures* Encyclopædia of Mathematics and its Applications, Vol. 67, Cambridge University Press, (1998)

[Ber08] O. Bernardi, *Solution to a combinatorial puzzle arising from Mayer's theory of cluster integrals* Seminaire Lotharingien de Combinatoire, 59, B59e (2008)

[BFP10] R. Bissacot, R. Fernández, A. Procacci, *On the convergence of cluster expansions for polymer gases* Journal of Statistical Physics 139, 598–617, (2010).

[BFPS11] R. Bissacot, R. Fernández, A. Procacci, and B. Scoppola, *An improvement of the Lovsz Local Lemma via cluster expansion* Combinatorics, Probability & Computing 20, 5, 709–719, (2011)

[BLG64] J. De Boer, J. M. J. Van Leeuwen, and J. Groeneveld *Calculation of the pair correlation function: II. Results for the 12-6 Lennard-Jones interaction* Physica 30, 12, 2265–2289, (1964)

[Bor01] É. Borel, *Leçons sur les séries divergentes. Vol. 2.* Gauthier-Villars, (1901)

[BoFu38] M. Born and K. Fuchs, *The statistical mechanics of condensing systems*, Proceedings of the Royal Society A 166, 391 (1938).

[BoGr49] M. Born, and H. S. Green *A general kinetic theory of liquids* Cambridge: University Press, (1949)

[Bou68] N. Bourbaki, *Elément de mathématiques: Théorie des ensembles* Chapter 4 'Structures', Hermann, Paris (1968)

[BoZa00] A. Bovier, and M. Zahradník, *A simple inductive approach to the problem of convergence of cluster exapsion in polymer models* Journal Statistical Physics, 100, 765, (2000)

[Bru83] N. G. de Bruijn, *The Lagrange-Good inversion formula and its application to integral equations*, Journal of Mathematical Analysis and Applications 92, 397–409 (1983)

[Bry84] D. C. Brydges, *A short course on cluster expansions*, Critical Phenomena, Random Systems, Gauge Theories. Les Houches Session XLIII, K. Osterwalder, R. Stora eds., Elsevier, 139–183, (1984)

[BrFe78] D. C. Brydges, P. Federbush *A new form of the Mayer expansion in classical statistical mechanics* Journal of Mathematical Physics 19, 2064 (1978)

[BrIm03] D. C. Brydges, and J. Z. Imbrie *Dimensional reduction formulas for branched polymer correlation functions* Journal of Statistical Physics 110, 3-6, 503–518, (2003)

[BrKe87] D. C. Brydges and T. Kennedy, *Mayer expansions and the Hamilton-Jacobi equation* Journal of Statistical Physics, 48, 19 (1987)

[BrMa14] D. Brydges, and D. H. U. Marchetti, *On the Virial Series for a Gas of Particles with Uniformly Repulsive Pairwise Interaction* arXiv preprint arXiv:1403.1621 (2014).

[Cai03] J. M. Caillol, *Some applications of the Lambert W function to classical statistical mechanics* Journal of Physics A: Mathematical and General, 36(42), 10431, (2003).

[Cam82] C. Cammarota, *Decay of Correlations for Infinite Range Interactions in Unbounded Spin Systems* Communications in Mathematical Physics, 85, 517–528 (1982)

[CJK97] R. M. Corless, D. J. Jeffrey, D. E. Knuth, *A sequence of series for the Lambert W function* ISSAC '97 Proceedings of the 1997 international symposium on symbolic and algebraic computation, 197-204 (1997)

[ClMc04] N. Clisby, and B. M. McCoy, *Negative virial coefficients and the dominance of loose packed diagrams for D-dimensional hard spheres* Journal of Statistical Physics 114, 5-6, 1361–1392 (2004)

[ClMc05] N. Clisby, and B. M. McCoy, *New results for virial coefficients of hard spheres in D dimensions* Pramana Journal of Physics, 64, 5, 775–783, (2005).

[ClMc06] N. Clisby, and B. M. McCoy, *Ninth and Tenth Order Virial Coefficients for Hard Spheres in D Dimensions* Journal of Statistical Physics, 122, 1, 15–57. (2006)

[Des13] M. Deserno, *van der Waals equation, the Maxwell construction and Legendre transforms* On-line notes: http://www.cmu.edu/biolphys/deserno/pdf/van-der-Waals-and-Maxwell.pdf (2013)

[Dob96] R. L. Dobrushin, *Estimates of semi-invariants for the Ising model at low temperatures* Translations of the American Mathematical Society-Series 2, 177, 59–82, (1996)

[DLL07] M. G. Ducharme, G. Labelle, P. Leroux, *Graph weights arising from Mayer's theory of cluster integrals* Séminaire Lotharingien de Combinatoire, 54, B54m, (2007).

[EhMe94] R. Ehrenborg, M. Mendez, *A bijective proof of infinite variated Good's inversion* Advances in Mathematics, 103 221–259 (1994)

[Far08] W. G. Faris, *A connected graph identity and convergence of cluster expansions*, Journal of Mathematical Physics, 49, (2008)

[Far10] W. G. Faris, *Combinatorics and cluster expansions*, Probability Surveys, 7, 157–206, (2010)

[Far11] W. G. Faris, *Combinatorial species and Feynman diagrams* Séminaire Lotharingien de Combinatoire, B61An, 37, (2011)

[Far12a] W. G. Faris, *The fundamental theorem of calculus for a matroid* Journal of Mathematical Physics, 53, 063305 (2012)

[Far12b] W. G. Faris, *Biconnected graphs and the multivariate virial expansion*, Markov Processes and Related Fields, 18, 357–386 (2012)

[FePr07] R. Fernández, and A. Procacci, , *Cluster expansion for abstract polymer models: new bounds from an old approach*, Communications in Mathematical Physics, 274, 1, 123-140, (2007)

[FePr08] R. Fernández and A. Procacci *Regions without complex zeros for chromatic polynomials on graphs with bounded degree* Journal of Combinatorics, Probability and Computing, 17, 225 (2008)

[Fey72] R. P. Feynman, *Lectures on Statistical Mechanics*, Addison-Wesley,(1972)

[FlSe09] P. Flajolet, R. Sedgewick, *Analytic Combinatorics* Cambridge University Press (2009)

[FNU56] G. W. Ford, R. Z. Norman, and G. E. Uhlenbeck, *Combinatorial problems in the theory of graphs. II*, Proceedings of the National Academy of Sciences of the United States of America, 42, 4, 203, (1956)

[FoUh56a] G. W. Ford, and G. E. Uhlenbeck, *Combinatorial problems in the theory of graphs. I*, Proceedings of the National Academy of Sciences of the United States of America, 42, 3, 122, (1956) .

[FoUh56b] G. W. Ford, and G. E. Uhlenbeck, *Combinatorial problems in the theory of graphs. III*, Proceedings of the National Academy of Sciences of the United States of America, 42, 8, 529, (1956)

[FoUh57] G. W. Ford, and G. E. Uhlenbeck, *Combinatorial problems in the theory of graphs. IV*, Proceedings of the National Academy of Sciences of the United States of America, 43, 1, 163, (1957)

[Fuc42] K. Fuchs, *The statistical mechanics of many component gases*, Proceedings of the Royal Society London A. 179, 408–432 (1942)

[Ges87] I. M. Gessel, *A combinatorial proof of the multivariable Lagrange inversion formula*, Journal of Combinatorial Theory 45, 178–195 (1987)

[GL95] I. M. Gessel, P. Leroux, *Lagrange inversion for species*, Journal of Combinatorial Theory, Series A 72, 1, 95–117 (1995)

[Good60] I. J. Good, *Generalizations to several variables of Lagrange's expansion, with applications to stochastic processes*, Proceedings Cambridge Philosophical Society, 56, 367–380 (1960)

[Good65] I. J. Good, *The generalization of Lagrange's expansion and the enumeration of trees*, Proceedings Cambridge Philosophical Society, 61, 499–517 (1965)

[Gor11] M. Gorzelańczyk. *Non-existence of phase transitions in continuous hard core gases with nonnegative potentials* Journal Statistical Mechanics: Theory and Experiment, 08, P08016, (2011)

[Gro62] J. Groeneveld, *Two theorems on classical many particle systems* Physics Letters, 3 50–51 (1962)

[Gro67] J. Groeneveld, *Estimation methods for Mayer's graphical expansions* in *Graph theory and theoretical physics* F. Harary, (ed.) Academic Press, (1967)

[GK71] C. Gruber, H. Kunz, *General properties of Polymer Systems* Communications in Mathematical Physics, 22, 133–161 (1971)

[HaPa73] F. Harary and E. Palmer, *Graphical Enumeration* Academic, New York, (1973)

[HaMa38] S. F. Harrison and J. E. Mayer, *The statistical mechanics of condensing systems. IV*, Journal of Chemical Physics, 6, 101 (1938)

[Hel14] T. Helmuth, *A Proof of the Brydges-Kennedy-Abdesselam-Rivasseau Formula by Kruskal's Algorithm* Private Communication (2014)

[HeLe70] D. Henderson and P. J. Leonard, *One- and two-fluid van der Waals theories of liquid mixtures, I. Hard sphere mixtures*, Proceedings National Academy of Sciences U.S.A. 67, 1818–1823 (1970)

[Hill56] T. L. Hill, *Statistical mechanics: Principles and selected applications* The McGraw-Hill Series in Advanced Chemistry, McGraw-Hill Book Co., Inc., New York (1956)

[Hus50] K. Husimi, *Note on Mayer's theory of cluster integrals*, Journal Chemical Physics, 18, 682, (1950)

[Isa94] S. B. Isakov, *Generalization of statistics for several species of identical particles* Modern Physics Letters B 8, 05, 319–327 (1994)

[IMO95] S. B. Isakov, S. Mashkevich, and S. Ouvry *On the thermodynamics of multispecies anyons* Nuclear Physics B 448, 3, 457–469 (1995)

[IsMa97] S. B. Isakov, and S. Mashkevich *Statistical mechanics and thermodynamics for multispecies exclusion statistics* Nuclear Physics B 504, 3, 701–718 (1997)

[Jan12] S. Jansen, *Mayer and virial series at low temperature* Journal of Statistical Physics, 147, 4, 678–706, (2012)

[JTTU14] S. Jansen, S. J. Tate, D. Tsagkarogiannis, D. Ueltschi, *Multispecies Virial Expansions* Communications in Mathematical Physics, 330, 2, 801–817 (2014)

[JaTs13] S. Jansen, D. Tsagkarogiannis, *Many-Species Tonks Gas* [In progress/Private Communication]

[Jon24] J. E. Jones, , *On the determination of Molecular Fields II - From the Equation of State of a Gas* Proceedings Royal Society London A, 106, 463 (1924)

[Joy81] A. Joyal, *Une théorie combinatoire des séries formelles* Advances in Mathematics, 42, 1, (1981)

[KaUh38] B. Kahn, and G. E. Uhlenbeck, *On the theory of condensation* Physica 5, 5, 399–416 (1938)

[Kam01] H. Kamerlingh-Onnes, *Expression of the Equation of State of gases and liquids by means of series* Communications from the Physical Laboratory of the University of Leiden 71, 3, (1901)

[Kam02] H. Kamerlingh-Onnes, *Expression of the Equation of State of gases and liquids by means of series* Proceedings of the Koninklijke Nederlandse Akademie van Wetenschappen, 4, 125 (1902)

[Kam64] N. G. van Kampen *Condensation of a Classical Gas with Long-Range Attraction* Physical Review, 135, A362, (1964)

[KaLe08] A. Kaouche, and P. Leroux, *Graph weights arising from Mayer and Ree-Hoover theories of virial expansions* DMTCS Proceedings 1 (2008).

[KiSa53] J. G. Kirkwood, Z. Salsburg, *The Statistical Mechanical Theory of Molecular Distribution Functions in Liquids* Discussions Faraday Society, 15, 28 (1953)

[KoRe96] A. Y. Kondratiev, and A. L. Rebenko, *Cluster expansions of Brydges-Federbush type for quantum lattice systems* Methods of Functional Analysis and Topology 2, 3, 59–68, (1996)

[Kot06] R. Kotecký, *Cluster expansions* Encyclopedia of Mathematical Physics 1, 531–536, eds. J. P. Françoise, G. L. Naber, and S. T. Tsou, Oxford: Elsevier, (2006)

[KoPr86] R. Kotecký, D. Preiss, *Cluster Expansion for Abstract Polymer Models* Communications in Mathematical Physics 103, 3, 491–498, (1986)

[Kra88] Ch. Krattenthaler, *Operator methods and Lagrange inversion: a unified approach to Lagrange formulas* Transactions of the American Mathematical Society 305, 2, 431–465, (1988)

[Lab81] G. Labelle, *Une nouvelle démonstration combinatoire des formules d'inversion de Lagrange* Advances in Mathematics 42, 3, 217–247, (1981)

[LLP84] L. D. Landau, E. M. Lifshitz, L. P. Pitaevskii, *Statistical Physics, Landau and Lifshitz Course of Theoretical Physics Vol. 5* Butterworth-Heinemann (1984)

[LePen64] J. L. Lebowitz, O. Penrose, *Convergence of virial expansions* Journal of Mathematical Physics, 5, 841–847 (1964)

[LePen66] J. L. Lebowitz, O. Penrose, *Rigorous Treatment of the Van Der Waals-Maxwell Theory of the Liquid-Vapor Transition* Journal of Mathematical Physics, 7, 98 (1966)

[LePer63] J. L. Lebowitz, J. K. Percus *Integral Equations and Inequalities in the Theory of Fluids* Journal of Mathematical Physics 4, 1495 (1963)

[LeRo64] J. L. Lebowitz and J. S. Rowlinson, *Thermodynamic properties of mixtures of hard spheres*, Journal of Chemical Physics 41, 133 (1964)

[LGB59] J. M. J. Van Leeuwen, J. Groeneveld, and J. De Boer *New method for the calculation of the pair correlation function. I* Physica 25, 7 792-808 (1959).

[LeDe37] J. E. Lennard-Jones, A. F. Devonshire, *Critical Phenomena in Gases I* Proceedings of the Royal Society London A, 163, 53 (1937)

[LeDe39] L. E. Lennard-Jones, and A. F. Devonshire *Critical and co-operative phenomena. III. A theory of melting and the structure of liquids* Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences 169, 938, 317–338 (1939)

[Ler04] P. Leroux, *Enumerative problems inspired by Mayer's theory of cluster integrals*, The Electronic Journal of Combinatorics, 11, 1, (2004)

[Lieb63] E. Lieb *New Method in the Theory of Imperfect Gases and Liquids* Journal of Mathematical Physics, 4, 671 (1963)

[LiPi80] E. M. Lifshitz, L. P. Pitaevskii, *Statistical Physics Part 2, Landau and Lifshitz Course of Theoretical Physics Vol. 9.* Pergamon Press, Oxford (1980)

[LPS98] B. N. B. de Lima, A. Procacci, and B. Scoppola, *A remark on high temperature polymer expansion for lattice systems with infinite range pair interactions* Letters in Mathematical Physics 45, 4, 303–322, (1998)

[Lyb05] I. Lyberg, *The fourth virial coefficient of a fluid of hard spheres in odd dimensions* Journal of Statistical Physics 119, 3, 747–764 (2005)

[Mac98] S. MacLane, *Categories for the Working Mathematician* Vol. 5. Springer Verlag (1998)

[MA37] J. E. Mayer and P. G. Ackermann, *The statistical mechanics of condensing systems. II*, Journal of Chemical Physics 5, 74 (1937)

[May37] J. E. Mayer, *The statistical mechanics of condensing systems. I*, Journal of Chemical Physics 5, 67 (1937)

[May39] J. E. Mayer, *Statistical mechanics of condensing systems V, two-component systems*, Journal of Physical Chemistry 43, 71–95 (1939)

[MMay40] J. E. Mayer, M. G. Mayer, *Statistical Mechanics* New York, John Wiley and Sons Inc. (1940)

[MMon41] J. E. Mayer E. Montroll, *Molecular Distribution* Journal of Chemical Physics, 9,2 (1941)

[May47] J. E. Mayer, *Integral Equations between Distribution Functions of Molecules* Journal of Chemical Physics 15, 187 (1947)

[MaSu91] A. E. Mazel, and Yu. M. Suhov *Random surfaces with two-sided constraints: an application of the theory of dominant ground states* Journal of Statistical Physics 64, 1-2, 111-134, (1991).

[McCoy10] B. McCoy, *Advanced Statistical Mechanics* Oxford University Press, (2010)

[MeNa93] M. Méndez and O. Nava, *Colored species, c-monoids, and plethysm. I*, Journal of Combinatorial Theory Series A 64, 102–129 (1993)

[MSV06] D. Merlini, R. Sprugnoli, M. C. Verri. *Lagrange Inversion: When and How* Acta Applicandæ Mathematica, 94, 3, 233–249 (2006)

[Mir00] S. Miracle-Solé, *On the convergence of cluster expansions* Physica A, 279, 244–249 (2000)

[MoPr13] T. Morais, A. Procacci, *Continuous Particles in the Canonical Ensemble as an Abstract Polymer Gas* Journal of Statistical Physics 151, 5, 830–849 (2013)

[Pen63a] O. Penrose *Convergence of Fugacity Expansions for Fluids and Lattice Gases* Journal of Mathematical Physics 4, 1312 (1963)

[Pen63b] O. Penrose *The Remainder in Mayer's Fugacity Series* Journal of Mathematical Physics, 4, 1488 (1963)

[Pen67] O. Penrose *Convergence of fugacity expansions for classical systems* in *Statistical mechanics: foundations and applications*, A. Bak (ed.), Benjamin, New York.

[PeYe58] J. K. Percus, and G. J. Yevick, *Analysis of Classical Statistical Mechanics by Means of Collective Coordinates* Physics Review 110, 1, (1958).

[Plu68] M. D. Plummer, *On minimal blocks* Transactions of the American Mathematical Society 134, 1, 85–94, (1968)

[Pro07] A. Procacci, *Abstract polymer models with general pair interactions*, Journal of Statistical Physics, 129, 1, 171–188, (2007)

[PoUe09] S. Poghosyan, and D. Ueltschi, *Abstract cluster expansion with applications to statistical mechanical systems*, Journal of Mathematical Physics, 50, 5, (2009)

[PT12] E. Pulvirenti, D. Tsagkarogiannis, *Cluster Expansion in the Canonical Ensemble* Communications in Mathematical Physics 316, 2 289–306, (2012)

[RH64a] F. H. Ree, and W. G. Hoover *Fifth and Sixth Virial Coefficients for Hard Spheres and Hard Disks* The Journal of Chemical Physics 40, 939 (1964).

[RH64b] F. H. Ree, and W. G. Hoover *Reformulation of the virial series for classical fluids* The Journal of Chemical Physics, 41, 1635 (1964)

[RH67] F. H. Ree, and W. G. Hoover *Seventh Virial Coefficients for Hard Spheres and Hard Disks* The Journal of Chemical Physics 46, 4181 (1967).

[RiTa14] V. Rivasseau, and A. Tanasa, *Generalized constructive tree weights* Journal of Mathematical Physics 55, 4, 043509, (2014).

[RiWa10a] V. Rivasseau, and Z. Wang *Loop vertex expansion for $\Phi^{2k}$ theory in zero dimension* Journal of Mathematical Physics 51, 092304, (2010).

[RiWa10b] V. Rivasseau, and Z. Wang *How are Feynman graphs resummed by the Loop Vertex Expansion?* arXiv preprint arXiv:1006.4617 (2010).

[Rota75] G. C. Rota, *Finite Operator Calclulus* Academic New York (1975)

[Rue63] D. Ruelle *Correlation Functions of Classical Gases* Annals of Physics, 25, 109 (1963)

[Rue64] D. Ruelle, *Cluster Property of the Correlation Functions of Classical gases* Reviews in Modern Physics 36, 2, 580–584, (1964)

[Rue69] D. Ruelle, *Statistical Mechanics* Benjamin, New York, (1969)

[Sim11] H. Simmons, *An Introduction to Category Theory* Cambridge University Press (2011)

[Sok01] A. D. Sokal, *Bounds on the complex zeros of (di) chromatic polynomials and Potts-model partition functions* Combinatorics, Probability & Computing 10, 1, 41-77, (2001)

[Sok05] A. D. Sokal, *The multivariate Tutte polynomial (alias Potts model) for graphs and matroids* Surveys in Combinatorics, 327, 173–226, (2005).

[Ste63] G. Stell *The Percus-Yevick equation for the radial distribution function of a fluid* Physica, 29, 517–534, (1963).

[Ste65] G. Stell, *Cluster Star Inversion by means of Generating Functionals* Journal of Mathematical Physics 6, 1193, (1965)

[Tar55] A. Tarski, *A Lattice-Theoretical Fixpoint Theorem and its Generalisations* Pacific Journal of Mathematics 5 282–309 (1955)

[Tate13] S. J. Tate, *Virial Expansion Bounds* Journal of Statistical Physics, 153, 2, 325–338 (2013)

[Tate14] S. J. Tate, *A Solution to the Combinatorial Puzzle of Mayer's Virial Expansion* arXiv:1402.2119

[Tem12] Ch. Temmel *Sufficient conditions for absolute convergence of hardcore cluster expansions* arXiv preprint:1209.4035v1 (2012)

[Thi85] M. Thiessen, *Untersuchungen über die Zustandsgleichung* Annalen der Physik 24, 467 (1885)

[Tut66] W. T. Tutte, *Connectivity in graphs* Volume 15. University of Toronto Press, (1966)

[Uel04] D. Ueltschi, *Cluster expansions and correlation functions* Moscow Mathematics Journal, 4, 511 (2004)

[Ush12] M. V. Ushcats, *Equation of State Beyond the Radius of Convergence of the Virial Expansion* Physical Review Letters 109, 4 040601 (2012)

[VaRa75] A. N. Vasil'ev, and R. A. Radzhabov, *Analysis of the nonstar graphs of the Legendre transform in the ising model* Theoretical and Mathematical Physics 23, 3, 575–579, (1975)

[Zei95] E. Zeidler, *Applied Functional Analysis* Volume 108, New York: Springer-Verlag, (1995)